Monitoring Animal Well-being

Gronskyte, Ruta

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Undesirable event detection in pig herd using video surveillance

Ruta Gronskyte

Kongens Lyngby 2015
PHD-2015-370

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads, building 324,
2800 Kongens Lyngby, Denmark
Phone +45 4525 3031
compute@compute.dtu.dk
www.compute.dtu.dk
PHD-2015-370
ISSN: 0909-3192

Summary (English)

In recent years, animal welfare in industrial slaughterhouses has become a significant concern for consumers, farmers, and meat producers. Different groups have different interpretations of animal welfare. For the majority of consumers, the definition of animal welfare is highly influenced by their values and experiences, whereas meat producers are interested in the stress animals endure and the effect this has on meat quality.

Pigs that arrive at slaughterhouses are more sensitive than usual for several reasons. Not all animals are accustomed to transportation and in some cases, they are transported over long distances. Upon arrival, it is common to mix pigs from different farms in one area. Such mixing can lead to fights between pigs, causing additional stress or harm. The unfamiliar environment can also increases their stress levels. Some industrial slaughterhouses handle up to 62,000 pigs per week. Ensuring the welfare of such a large number of pigs using only personnel is a complicated task.

Video surveillance of humans has been widely used to ensure safety and order in various situations. Methods have been developed to detect individual actions or abnormal behavior in small groups and dense crowds. In recent years, surveillance has also been used to monitor animals. Research has mainly focused on monitoring laboratory animals and farm animals. In both cases, animals are usually in a constrained environment and cameras cover all areas where animals are present. Non-intrusive markers or extracted features are used for tracking in order to obtain better results. Laboratory environments can be highly controlled and therefore no light or shadow noise are present in videos.

In slaughterhouses, the main focus is on monitoring large groups of animals in locations where additional markers cannot be used and pigs can leave or enter

the surveyed area. In addition, pigs have a specific walking pattern, making motion analysis challenging. The first aim of this thesis was to monitor the movement of pigs without using any additional markers or feature extraction in an unconstrained environment.

In video surveillance, the behavior of humans and animals is monitored based on extremes: an event is present/event is not present, objects behave normally/objects behave abnormally, action 1/action 2/action 3, and so on. The motion of humans and animals is continual with transitions from one action to another. The second aim of this thesis was to propose a method to monitor motion as a continuous process using common classification methods.

In this thesis, monitoring was performed by employing optical flow (OF) using color images. This approach has not previously been used in this context. The color images provide better results than thermal images. An OF-based approach together with low-level statistics has previously been used to monitor chicken welfare. However, this approach is not suited to pigs, and instead modified angular histograms were used. In addition, a framework for continuous motion monitoring, using the principles of multivariate statistics and statistical process control is presented. The proposed framework increases event detection by 25% when compared to discrete motion monitoring. A lens correction method using a moving reference object is also proposed. Three tests indicated that the method performed well. This thesis presents promising ideas for a new approach for assignable cause identification in a highly auto-correlated process.

Summary (Danish)

I de seneste år er dyrs trivsel i industrielle slagterier blevet en væsentlig bekymring for forbrugere, landmænd og kødproducenter. De tre segmenter har forskellige fortolkninger af dyrs trivsel. For de fleste forbrugere er dyrs trivsel stærkt påvirket af deres værdier og erfaringer. Kødproducenter er interesseret i stress dyrene kan holde til, fordi det påvirker kødkvaliteten.

Grise der ankommer til slagterierne er mere følsomme end normalt af flere grunde. I nogle tilfælde transporteres grisene over lange afstande. Ikke alle dyr er vandt til transport. Ved deres ankomst, er det almindeligt at grise fra forskellige landmænd samles i et område. Den blanding kan forårsage kampe mellem svin, hvilket kan føre til yderligere stress eller skader. Det ukendt miljø øger også deres stressniveau. I nogle industrielle slagterier håndteres der op til 62.000 svin om ugen. Sikring af velvære af sådanne store antal svin udelukkende ved brug af personale er en kompliceret opgave.

Videoovervågning af mennesker har været meget anvendt til at sikre sikkerhed og orden i flere situationer. En række metoder er blevet udviklet til at fange individuelle handlinger eller unormal adfærd i små grupper og tætte folkemængder. I de senere år er overvågning også blevet anvendt til dyr. Forskningen har hovedsagelig fokuseret på overvågning af forsøgsdyr og husdyr. I begge tilfælde er dyrene i begrænsede miljøer og kameraer dækker alle områderne hvor der er dyr. For at opnå bedre resultater bruges ikke-invasive markører eller særlige egenskaber til sporing. Laboratoriemiljøer er meget kontrollerede; intet lys eller skygger i videoerne.

På slagterier det primære fokus at overvåge store grupper af dyr på steder, hvor der ikke kan bruges yderligere markører. Grisene kan også forlade eller komme ind i det overvågede område. Desuden har grise et specifikt gangmønster, hvor-

for bevægelsesanalyse ikke er ligetil. Det første formål med denne afhandling er at overvåge grises bevægelse uden brug af yderligere markører i et ubegrænset miljø.

I videoovervågning er adfærden hos mennesker og dyr overvåget på grundlag af ekstremer: en begivenhed er enten til stede eller ej, objekter opfører sig normalt eller ej, handling 1, 2, eller 3, osv. I naturen er menneskers og dyrs bevægelse kontinuerlig med overgange fra en handling til en anden. Det andet formål med denne afhandling er at foreslå en metode til overvågning af bevægelser som en kontinuerlig proces ved hjælp af standard klassificeringsmetoder.

I denne afhandling blev overvågningen udført med optisk bevægelse (optical flow, OF) på farvebilleder. Denne tilgang er ikke tidligere brugt i denne sammenhæng. Farvebillederne giver bedre resultater sammenlignet med termiske billeder. Brugen af OF og basal statistik er tidligere brugt til at overvåge kyllingevelfærd. Dog er den tilgang ikke brugbar på grise, hvorfor modificerede vinkelhistogrammer blev brugt. Dertil er en metode for overvågning af kontinuerlig bevægelse ved brug af principper fra multivariat statistik og statistisk processkontrol også præsenteret. Den viste metode øger detektionen af hændelser med 25% sammenlignet med en diskret bevægelsesovervågning. En metode til linsekorrektion med et ikke-stationært referenceobjekt er præsenteret. Tre test indikerer at metoden virker. Denne afhandling præsenterer lovende idéer til en ny tilgang til at identificere årsagssammenhænge i meget korrelerede processer.

Preface

This thesis was prepared at the Technical University of Denmark (DTU), Department of Applied Mathematics and Computer Science (DTU Compute), Section of Statistics and Data Analysis in partial fulfillment of the requirements for acquiring a Ph.D. degree in Applied Mathematical Statistics. The project was funded by the Technical University of Denmark and the Danish Meat Research Institute. The project was supervised by Murat Kulahci, Line Katrine Harder Clemmensen, and Marchen Sonja Hviid.

The thesis focuses on monitoring animal well-being using video recordings. It consists of four research papers.

Lyngby, 04-May-2015

Ruta Gronskyte

List of contribution

Manuscripts

A **R. Gronskyte**, M. Kulahci, L. K. H.Clemmensen, "Monitoring Motion of Pigs in Thermal Videos", in *Workshop on Farm Animal and Food Quality Imaging*, pp. 31-36, 2013.

B **R. Gronskyte**, L. H. Clemmensen, M. S. Hviid, M. Kulahci, "Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms", *Biosystems Engineering*, vol. 141, pp. 19-30, 2016.

C **R. Gronskyte**, L. H. Clemmensen, M. S. Hviid, M. Kulahci, "Pig herd monitoring and undesirable tripping and stepping prevention", *Computers and Electronics in Agriculture*, vol. 119, pp. 51-60, 2016.

D **R. Gronskyte**, L. Clemmensen, M. Kulahci, "Online Monitoring of Crowds", Submitted to: Computer Vision and Image Understanding Journal.

Presentations

Invited

- "Behavior analysis using optical flow", *Visionday 2013*, May 22, 2013, Copenhagen, Denmark

Contributed

- Presented a paper: **R. Gronskyte**, M. Kulahci, L. K. H.Clemmensen, "Monitoring Motion of Pigs in Thermal Videos", in *Workshop on Farm Animal and Food Quality Imaging*, pp. 31-36, 2013, Espoo, Finland.
- Presented a paper: **R. Gronskyte**, L. Clemmensen, M. Hviid, M. Kulahci, "Monitoring Motion of Pigs", *Farm Animal Imaging*, pp. 113-115, September 25-26, 2014, Taarstrup, Denmark.

Posters

- Presented a paper: **R. Gronskyte**, L. Clemmensen, M. Hviid, M. Kulahci, "Monitoring Motion of Pigs", *International Computer Vision Summer School*, July 14-20, 2013, Calabria, Italy.
- Presented a paper: **R. Gronskyte**, L. Clemmensen, M. Hviid, M. Kulahci, "Monitoring Motion of Pigs", *Farm Animal Imaging*, September 25-26, 2014, Taarstrup, Denmark.

Other contributions

This report is not part of the methodological developments included in this project and will not be further addressed:

- T. Sliusarenko, **R. Gronskyte**, L.K.H. Clemmensen, M. Kulahci, B.K Ersbøll,"Study Delay", *DTU Compute-Technical Report-2014*, 2014.

Acknowledgements

I would like to express my appreciation and thanks for the tremendous help from my supervisors, Murat Kulahci and Line Katrine Harder Clemmensen at the Technical University of Denmark. I would like to thank you for encouraging me in my research and for all your advice, which helped me navigate the science world. Your insightful comments were priceless.

A special thank you goes to my external supervisor, Marchen Sonja Hviid from the Danish Meat Research Institute. You helped me learn about pigs and their behavior and you provided me with data.

Thank you to Pia Brandt for all the time you spent explaining pigs and collecting data from the Danish Meat Research Institute. I would also like to thank Pia's supervisor, Margit Aaslyng.

I would like to thank Fugee Tsung for hosting me at the Hong Kong University of Science and Technology. I am thankful for your comments on my research and for the many invitations to scientific activities.

I would like to give a special thank you to my husband Rune for all your moral support and long discussions about my research.

Thank you to my mom, dad, and sister for your support and encouragement.

Thank you to my dear friend Laura for always taking time to listen.

Abbreviations

ARL	*Average run-length*
ARMA	*Autoregressive Moving Average*
CDV	*Classification Decision Value*
CUSUM	*Cumulative Sum*
EWMA	*Exponential Weighted Moving Average*
IA	*Image Analysis*
LCL	*Lower Control Limit*
LOESS	*Locally Scatterplot Smoothing*
LOWESS	*Locally Weighted Scatterplot Smoothing*
MAH	*Modified Angular Histogram*
MC	*Markov Chains*
MS	*Multivariate Statistics*
OF	*Optical Flow*
PCA	*Principal Component Analysis*
PC	*Principal Component*
RL	*Run-Length*
SPC	*Statistical Process Control*

SVM	*Support Vector Machines*
UCL	*Upper Control Limit*

Symbols

(x, y) and (a, b)	2D point coordinates
(x, y, z)	3D point coordinates
t	Time
$I(x, y, t)$	Image intensity at point (x, y) and time t
$\mathbf{x}$	Vector of features
$(u; v)$	Optical flow vector components
Ω	Image area
λ	Weight parameter
B	Structuring element
E	Shape in a image
$\delta_B(E)$	Shape dilation function
(x_D, y_D)	Distorted image points
(x_c, y_c)	Distortion center
κ_{D1}, κ_{D2}	Lens distortion coefficients
p_1, p_2	Tangential distortion coefficients
s_x, s_y	Frame size
l_x, l_y, l_z	Surveilled area size
k_z	Distance from 3D point to the camera center
X	Data matrix
T	Scores of principal components
W	Loadings of principal components
$G(\mathbf{x})$	SVM classification function
C	Constant
d	Degree of the polinomial
$K_\lambda(x_0, x_i)$	Weight function
$f(x)$	Function of x

β, β_0	Linear function coefficients
$K(\mathbf{x}, \mathbf{x}')$	Kernel function
ξ_i	Slack variable
κ_1, κ_2	Coefficients of neural network basis function
γ	Coefficient of radial basis function
μ_w	Mean of the process
σ_w	Standard deviation of the process
L_{cl}	Size of the control limits
Pr	Probability mass function
s	Number of independent Bernoulli trials
n	Sample size
p, q	Probabilities
$\bar{\bar{x}}$	Mean of a sample means
R	Sample range
A_2, D_3, D_4, B_3, B_4	A factor of control limits
S	Covariance matrix
i	Index
M	Target value of the process
K	Control limits
C_i^-, C_i^+	CUSUM chart control limits
D	Number of nonconforming units that are expected in the process
m	Number of samples
Δ	Shift difference
z	Random variable
c	Poisson distribution mean and variance
$\bar{u}$	Average number of nonconformities per unit
ϕ_i	Coefficients of auto-regressive model
θ_j	Coefficients of moving average model
ε_t	White noise
ϵ	Error between predicted and observed characteristics
T^2	Hotelling statistics
L	Number of batches
J	Number of characteristics observed
K	Time span of observations
ϑ	Transition probabilities
τ	Stationary distribution
Γ	Transition matrix
U	Eigenvectors
w	Eigenvalues
$Z(x, y)$	Polynomial function
m_l, m_w	Length and width measurements
$[o; r]$	Re-scaling interval

k	Number of bins
g_1	Skewness
σ_{g_1}	Standard deviation of a skewness
N	Smoothing window size
Q	Number of principal components
ψ	A probability of not detecting a shift on the first subsequent sample

Contents

Chapter 1

Introduction

This thesis considers automatic description and detection of animal welfare at slaughterhouses. The concept of animal welfare is first described in this chapter, which is followed by the aim of this project, the contribution of this thesis and the data used for the analysis.

1.1 Animal welfare

Animal welfare is a major concern for many consumers, farmers, and meat producers. Different aspects of animal welfare are of interest to different groups [1]. In recent years, there has a been growing concern among consumers regarding not only the quality and price of meat, but also criteria such as animal welfare and the environmental effects of livestock production. These criteria are influenced by the values and emotional experiences of the individual [2]. Meat producers are interested in the aspects of animal welfare that can affect meat quality and price.

D. M. Broom (1986) [3] defines welfare as the ability of an individual to, *"cope with its environment. Coping can sometimes be achieved with little effort and expenditure of resources, in which case the individual's welfare is satisfactory."* Specific effects of a failure to cope with the environment include increased rates

of mortality and disease, in addition to a decrease in size and the number of offspring. Both psychological and physical stress must be overcome in order for an individual to cope with their environment [4]. Psychological stress can be caused by changes in surroundings or inadequate handling, while unsuitable climate or inadequate supplies of oxygen, water, and food can lead to physical stress. Farmers (as well as meat producers) are interested in the aspects of animal welfare that can affect the health of the livestock and the meat quality.

The inappropriate handling of pigs during transportation both to and within slaughterhouses increases the levels of glucose, lactate, and other components in the blood [5, 6] before slaughter, which can affect meat quality. Furthermore, transportation (including loading and unloading of animals on and from vehicles), can cause notable injury [7]. Stressful events, such as animals tripping over and stepping on each other, are common during unloading, and can cause visible injuries such as bruises and scratches, which can affect the meat quality. After unloading, pigs that are unfamiliar with each other may start fighting, causing additional psychological or physical stress.

In some industrial slaughterhouses, an average of 62,000 pigs are handled every week. Tracking and ensuring a stress-free environment for such a large number of animals using only human labor is difficult. Visual surveillance is very useful in automating labor-intensive tasks and conducting detailed behavior monitoring. A. R. Frost et al. (1997) [8] suggest creating an integrative system, in which information about animals is collected in various ways to ensure animal welfare. In this thesis, a methodology for pig behavior analysis in slaughterhouses is presented, which can be used to create such an integrative system.

1.2 Project aim

The research presented in this thesis focuses on monitoring pigs in a slaughterhouse. The aim of this thesis is to provide a tool that can help ensure animal welfare by detecting *undesirable events*, (referred to in this thesis as *events*) such as animals tripping over and stepping on each other. Such events most commonly occur when animals are unloaded from transportation vehicles at a slaughterhouse. In this study, animal unloading was recorded using a GoPro HERO2 (©2013 Woodman Labs, Inc.) camera, generating the data set *unloading*. Two additional data sets were used for pig movement analysis: *entrance*, recorded using the same camera, and *thermal*, recorded using a thermal camera.

Several methodological hurdles had to be overcome to achieve the aims of this thesis. The first hurdle was related to monitoring a herd in an unconstrained environment. Several publications (as reviewed in the subsequent sections of this thesis) discuss tracking and monitoring individual animals, but in a con-

strained environment. The second hurdle concerned frame distortions due to the camera position. The third hurdle pertained to the general surveillance framework. It is common to use discrete classification in detecting abnormal behavior. However, as will be explained later, motion is continuous. To overcome these challenges, methodologies of image analysis, multivariate statistics, and statistical process control were combined. The solution allowed continuous monitoring of animals using video recordings in a binary classification problem where multiple variables must be analyzed.

1.3 Contributions

This thesis combines three academic disciplines: computer vision, multivariate statistics, and statistical process control. Some of the methods and principles used in this thesis are well known, though to the best of our knowledge, have not previously been used in conjunction with other disciplines in the context presented here. For example, the final crowd behavior classification is achieved using principles of statistical process control, where continuous decision values of support vector machines are used in conjunction with a cumulative sum chart. Other methods such as lens correction and modified angular histograms are proposed in this thesis. Since the main application aims to provide a procedure for the detection of undesirable animal behavior, it is expected that it will be of greatest interest for practitioners from agriculture and animal husbandry. Academics and computer vision specialists may also be interested in the proposed ideas and methods of statistical process control in decision-making.

1.4 Data

This project uses three data sets. The first data set includes grayscale thermal videos and is named *thermal data set*. The second and third data sets include color videos recorded at the same slaughterhouse. The first color data set was recorded while pigs were unloaded at the slaughterhouse and is named *unloading data set*. The second color data set was recorded after unloading, when the pigs had just entered the slaughterhouse, thus this data set is named *entrance data set*. In the following subsections, more details of the three data sets are given.

1.4.1 Thermal data set

Originally, the *thermal data set* was recorded for a pig counting project. Some pre-processing was performed during the original project to remove noise created by dirt and markings on the pigs. The sample frame is presented in Figure 1.1.

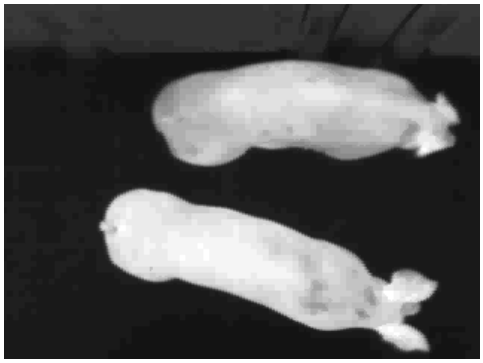


Figure 1.1: Example of the thermal data set.

In the video, the pigs move from left to right in a corridor. The camera view encompasses the entire width of the surveyed corridor. This data set was considered the starting point for this thesis. The initial expectation was that the use of thermal videos would allow the easy identification of pigs in a frame because pigs are much lighter in color than the background. However, since thermal cameras are restrictively expensive, this idea was not considered further.

1.4.2 Unloading data set

The general slaughterhouse entrance plan for the surveyed area is presented in Figure 1.2. "Camera 1" was used to record the unloading process. The camera was located above the entrance, pointing towards a truck. This camera position resulted in foreshortening distortion, where pigs further away from the camera appeared smaller than those closer to the camera. This video recording was also subject to lens distortion, where objects closer to the optical axis of the camera appear magnified. A detailed description of lens and foreshortening distortions can be found in 2.3. A sample frame from the *unloading data set* is presented in Figure 1.3.

The pigs in the truck were divided into pens, with approximately 20 pigs in each pen. In the video, the unloading of one pen is referred to as a "clip," and the part

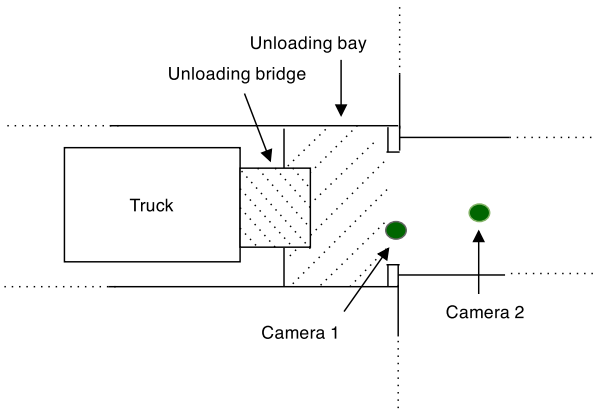


Figure 1.2: Slaughterhouse plan of the surveyed areas. Green dots represent camera positions.

Figure 1.3: Example of the unloading data set.

of the clip that is annotated is named "section". During transportation, pigs are sedentary and are therefore slow and unwilling to move during unloading. The personnel use specially designed sticks that make sounds to expedite the unloading. Some pigs can start moving too fast, resulting in events such as tripping over or stepping on each other. The *unloading data set* is used to analyze events in Papers C and D.

1.4.3 Entrance data set

"Camera 2" in Figure 1.2 was used to record the *entrance data set*. The camera was positioned perpendicular to the surveyed area, thus only lens distortion was present in the videos. A sample frame from the *entrance data set* is presented in Figure 1.4. Very few undesirable events were present in this data set. When pigs are encouraged to move, they will move from bottom to top of a frame. If they are not encouraged to move in a certain way, they will make random movements, lie down, or interact with each other. This data set is mainly used to understand the pigs' movement; the results are presented in Paper B.

Figure 1.4: Example of the entrance data set.

1.5 About this thesis

This thesis combines three main areas of statistics: *Image Analysis* (IA), *Multivariate Statistics* (MS), and *Statistical Process Control* (SPC). Readers are not expected to be familiar with all areas and therefore a brief introduction to each topic is presented in Chapters 2, 3, 4, and 5. A literature review and references are provided for further in-depth self-study. These introductions can be read in any order or can be disregarded if the reader is familiar with the topic. The

remainder of the thesis will introduce new concepts and merge the three methodological areas; this material should be read in chronological order. Chapter 6 presents a proposed framework for monitoring pigs' behavior and the methods for each step. The methods are presented in the same order that they are used in the proposed monitoring approach. The results are presented in Chapter 7, and a discussion is presented in Chapter 8, followed by a conclusion.

Chapter 2

Motion estimation and object detection

As mentioned in Chapter 1, several steps are involved in monitoring animal behavior in video recordings. An overview of the image analysis steps is presented in this chapter.

2.1 Optical Flow

Motion of objects in video recordings can be captured using two main approaches: (1) by identifying an object in each frame and track the coordinates throughtout the recording; or (2) using *Optical Flow* (OF) ([9, pp. 239-258],[10, pp. 5-34]), which estimates the relative motion for every pixel in a frame. The second approach is used in this thesis because it is the most popular choice in the literature on crowd monitoring. Additional reasons behind its use will be discussed later in this chapter.

OF estimation has attracted a lot attention since the concept was first published in [11]. Several methods for OF estimation have been proposed in the literature, which are based on either gradient, phase correlation, and spatio-temporal energy and others. The most popular methods are gradient based,

such as Lucas-Kande [12] and Horn-Schunk [13]. The latter is used to estimate OF in this thesis and is presented further is this chapter.

For the two-dimensional case, the grayscale pixel intensity at location (x, y) and time t is noted as $I(x, y, t)$. Assume that the intensity of pixel $\mathbf{x} = (x, y)$ does not change over time. The movement of $\mathbf{x}$ from time t to time $t + 1$ is defined as:

$$I(a, b, t) = I(x + u, y + v, t + 1) \tag{2.1}$$

where (u, v) represents the components in the OF vector and (a, b) is a new location of pixel intensity. If the movement is small, the first-order Taylor approximation can be used:

$$I(x, y, t) \approx I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt \tag{2.2}$$

I_x, I_y and I_t denote the partial derivatives. To estimate (u, v), we have to solve:

$$0 = I_t + I_x u + I_y v \tag{2.3}$$

which does not have a unique solution. The constant intensity assumption is violated in practice due to changing illumination and the camera noise. One more complication known as the *aperture problem* arises when the motion of un-textured objects is of interest. To overcome these complications smoothness and regularization constraints are imposed during the estimation process. The regularization methods are based on (1) features of the image or (2) variations in the brightness intensities of the pixels. The former estimates motion for each pixel individually, while the latter considers information about neighboring pixels and has smoothing properties. This is exactly the Lucas-Kanade and Horn-Schunk methods, respectively.

Typically the variation regularized methods give better results, thus only the Horn-Schunk approach will be considered. This method solves the minimization problem:

$$\min_{u(\mathbf{x}),v(\mathbf{x})} \left\{ \int_\Omega (|\nabla u(\mathbf{x})|^2 + |\nabla v(\mathbf{x})|^2) d\Omega + \lambda \int_\Omega (I_t + I_x u(\mathbf{x}) + I_y v(\mathbf{x}))^2 d\Omega \right\} \tag{2.4}$$

where Ω is an image area and the vector components of pixel $\mathbf{x} \in \mathbb{R}$ displacement are defined as $u \in \mathbb{R}$ and $v \in \mathbb{R}$. The first integral in Equation 2.4 is a high-variation penalty, the second integral is a quadratic constraint, and λ is a weight parameter.

2.2 Optical flow filtering

OF is estimated for an entire image; therefore, filtering is employed to identify only those OF vectors that belong to pigs. One of the most basic approaches used in practice is based on the assumption that the background does not move; therefore, thresholding vectors that are too short can be applied. However, setting such thresholds is not possible in the analyzed case because pigs' movements can be very diverse. Some pigs might lie down and not move at all, while others might run. The former behavior causes most of the difficulties. The OF vectors of stationary pigs will be approximately the same length as other stationary objects in the frame.

Two other approaches are applied to filter OF in this thesis: (1) blob detection and (2) background subtraction with blob detection. The former detects regions with common properties. The latter approach subtracts the image without pigs from the current image and then identifies blobs in the remaining image. The background subtraction with blob detection method is often preferred because it is more robust. Median filtering, hole filling, and object dilation are used to correct minor errors around and in the detected pigs.

2.2.1 Methodology of blob detection

Blob detection is a useful tool for segmenting images into regions with common properties. Identified blobs can be used for tracking, texture analysis, object recognition, and other feature analysis. The most primitive approach is based on thresholding the pixel color values, but this method is very sensitive to camera and illumination noise. A large number of methods can be used for blob detection; it is not possible to discuss all of them in this thesis. In this section, a rough classification of the methods is presented, and the method used for analysis is presented in greater detail.

Blob detection methods can be classified in two groups based on (1) derivatives and (2) morphological operations. Hessian and Harris operator-based methods use second- and first-order derivatives of image intensity functions, respectively. The Laplace operator and affine shape adaptation are applied to achieve more robust results.

Methods based on morphological operations use information about the local maximum (minimum) to identify blobs in a grayscale image. This is achieved by viewing images as topographic surfaces. The gravitation law indicates that water on the topographical surface will end up in a basin. By continuing to fill the surface with water, the watersheds will divide the surface into zones, with

each containing a basin. This approach is known as watershed transformation [14, p. 267-292]. Watershed transformation is not suitable for cases in which plateaus are present on the surface. The alternative approach is based on a flooding concept. Holes are made in each basin, and the surface is submerged in water. The water will start filling the basins, and dams are installed where water in each basin begins to connect. At the end of the process, dams represent the watersheds, and each watershed is labeled as a separate region. This flooding approach is implemented in Matlab [15]. The algorithm repeats two steps until the image is divided into the regions:

1 Select next unlabeled pixel i.

2 Using flood algorithm, label all the pixels in connection to i.

Numerous other methods can be used for blob detection, each of which has different properties. The choice of method usually depends on the complexity of the images and the available computational time.

2.2.2 Other morphological operations

The theory of morphological operations in computer vision is not just used for blob detection. It is a powerful tool for image manipulation because it analyzes the shapes and forms of objects. Three other functions based on morphological operations are used in this thesis: median filtering, hole filling, and dilation [14, p. 55, 208, 65-70]. Median filtering can be used to remove noise in an image. This method works by replacing a pixel's color value with the median color value of neighboring pixels. A hole in a grayscale image is an area with minimal quantities that are not at the edge of an image. The hole-filling method removes areas with minimal quantities. The structuring element must be presented first to explain the dilation method. A structuring element is a shape (e.g., disk, diamond, or square) that is used to analyze or modify a given image. If B is a structuring element and E is a shape in an image that will be dilated, then shape dilation function is:

$$\delta_B(E) = \{e|B_e \cap E \neq \oslash\} \tag{2.5}$$

In this thesis, the morphological operations described above are used in the pig identification step in a frame.

2.2.3 Methodology and literature review of background subtraction

One way to identify moving objects in a frame is to use background subtraction. The *background* of an image is a monitored scene without moving objects, while the *foreground* is a frame with only moving objects. All background subtraction methods have three common steps: (1) establish the background frame; (2) subtract the background frame from the current frame, and (3) remove the noise and label the objects. A large number of background subtraction methods have been proposed in the literature. The choice of method depends on the overall complexity of the scene and the available computation time. [16] considers the following challenges when comparing background subtraction methods: gradual illumination and sudden illumination changes; dynamic backgrounds (some parts of the background can be of a dynamic nature); camouflage (background and foreground objects are similar in color); shadows; bootstrapping (background image is not available without the foreground image); and camera noise. A comparison of several background subtraction methods based on speed, memory, and accuracy can be found in [17].

The background subtraction method used in this thesis is based on the **HSV** (hue-saturation-value) color system and a bit-wise comparison. In the first step, images are converted from the **RBG** to the **HSV** color system. **HSV** is a cylindrical color representation, in which **H** represents a color attribute, **S** represents the ratio of colorfulness to brightness, and **V** represents the ratio of brightness to a similarly illuminated white. [18] suggests that **HSV** color representation is a useful tool in eliminating shadows and bright areas in an image. In the second step, a fast bit-wise comparison *XOR* is performed between the current frame and the background frame. The comparison is performed for each element of the pixel vector. If both elements are 0 or 1, it returns 0; otherwise, it returns 1. The accuracy of the presented background subtraction method is compared to the blob detection method in Section 7.2.

2.3 Lens and foreshortening distortions

It is common to use a radial lens distortion that allows survey a larger area of interest. A perspective projection, which creates a foreshortening distortion, is a less common but possible distortion used to cover a larger area of interest. The correction of individual distortions is widely discussed in the literature, but the combination of distortions is not. Little research has been done to evaluate the effects on estimated motion from distorted videos. [19] combines lens distortion correction and OF estimation. In practice, if the lens parameters are known,

correction can be performed.

In pig monitoring cases, both types of distortions are present and neither the lens parameters nor the camera position is known. Radial lens distortion has two sub-types: (1) pincushion distortion and (2) barrel distortion. The former creates an effect in which lines slant inwards although they are straight in real life. The latter creates an effect in which lines bend outwards. Radial distortion can be defined as:

$$\begin{aligned} x_D &= x(1 + \kappa_{D1} g^2 + \kappa_{D2} g^4) + 2p_1 xy + p_2(g^2 + 2x^2) \\ y_D &= y(1 + \kappa_{D1} g^2 + \kappa_{D2} g^4) + 2p_2 xy + p_1(g^2 + 2y^2) \end{aligned} \tag{2.6}$$

where $g = \sqrt{(x - x_c)^2 + (y - y_c)^2}$, x_D, y_D; and x, y are distorted and undistorted image points, respectively; x_c, y_c distortion center; κ_{D1} and κ_{D2} are lens distortion coefficients; and p_1, p_2 are tangential distortion coefficients, which can be caused when the lens is not perfectly aligned.

An object seen in a 3D-perspective projection can be remapped in 2D using the following formula:

$$\begin{aligned} x_D &= l_x \frac{x s_x}{k_z l_x} \\ y_D &= l_y \frac{y s_y}{k_z l_y} \end{aligned} \tag{2.7}$$

where s_x, s_y and l_x, l_y are the size of the frame and the surveilled area, respectively; l_z is the distance from the surveilled area to the center of camera; k_z is the distance from the 3D point to the camera center; and x, y is a point in 3D space.

A new method is proposed in Paper C and Section 6.2 providing a correction method for combination of lens and foreshortening distortions. The method is fast and simple to implement. Unfortunately, the application of this method is narrow because it is common to use more cameras to account for distortion.

2.4 Surveillance

Surveillance has received extensive attention due to its many possible applications. It is used to ensure public safety at special events, identify people, manage crowds, and detect abnormalities. In recent years, surveillance has also been used to ensure animal welfare and detect health-related events. Automated monitoring is not labor intensive, and it allows people to analyze multiple areas simultaneously. Also, accurate positions and times of events can be tracked.

A large number of methods have been developed to accommodate different situations and events of interest. In this section, an overview of the surveillance concept is presented. A flowchart of a general framework for visual surveillance using a single camera is presented in Figure 2.1.

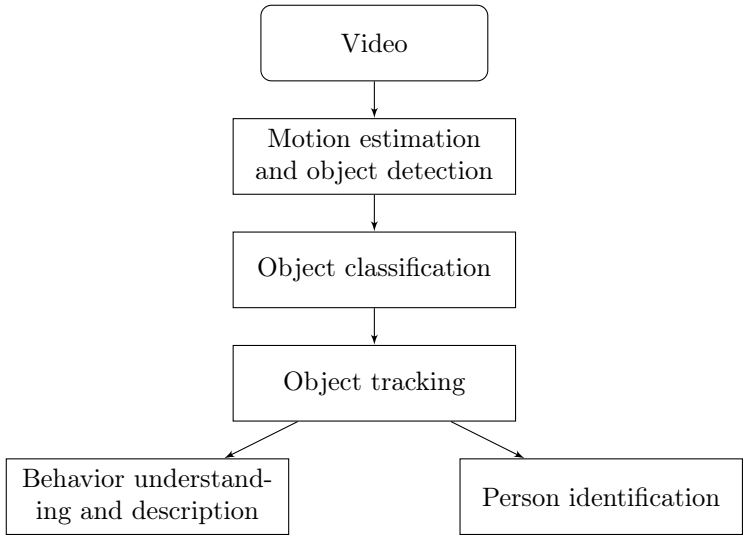


Figure 2.1: General surveillance framework using a single camera.

Information from several cameras can be combined at any step in the framework. For each step, a large number of methods are proposed. The choice of method depends on general objectives, surveyed objects, lighting, and camera position, among other factors. In most cases, the background of the image must be identified to estimate motion and detect objects of interest. In some cases, the entire crowd is of interest. In other cases, individual objects are of interest. Object tracking is one of the most researched areas in surveillance, and it is highly challenging. The method used must be able to handle occlusions and interacting objects [20]. The patterns of moving objects can be used to design public places, ensure safety, and track persons of interest. An in-depth overview of available methods can be found in [21, 22, 23].

One special use for surveillance is in crowd monitoring. Large, dense crowds are common at music events, protests, and religious gatherings. The are two approaches to analyze motion that are based on: (1) non computer vision and (2) computer vision methods. The approach that is not based on computer vision uses principles of psychology and sociology. It often includes other information that is not directly available in images, such as planning of an area, prior knowledge of human interactions, and group behavior. A computer vi-

sion approach extracts features that are sometimes modeled and then used for classification. The three most-used features, for dense crowd analysis, are (1) density, (2) speed, and (3) location. Individual people are rarely identified and tracked in dense crowds due to complications in separating objects and problems that arise when objects occlude or interact. In the review paper [24] it is stated that individual object-tracking techniques are only reasonable for up to ten people. In this thesis, computer vision methods are applied because no social and psychological studies have been conducted on the movement of pig herds. The most common technique for estimating motion in computer vision approach is based on OF. The framework is presented in Figure 2.2.

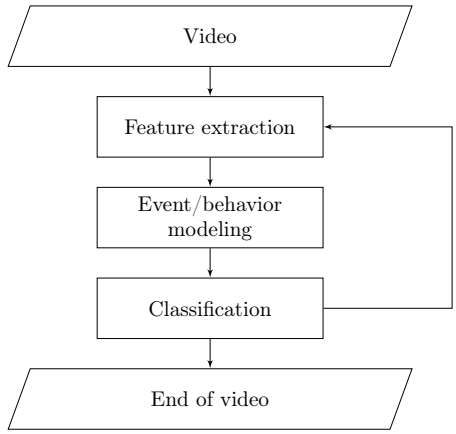


Figure 2.2: Dense crowd monitoring flowchart.

In computer vision approach it is common directly threshold ([25, 26, 27, 28]) features ([29, 30, 27, 28, 31]). The hidden Markov model ([25, 32]), social force model [33], social entropy measure [34], or a mixture of dynamic textures [26] are also used to model a crowd. The most commonly used classification methods are the *Support Vector Machines* (SVM) ([34, 29, 30]), K-mean [33], and neural network [31] methods.

Animal behavior monitoring is used mainly in two areas: laboratories and industrial farms. The aim of animal monitoring in laboratories is to track individual animals and measure the length and frequency of their activity. To ensure that animals are not disturbed by surroundings, the environment is highly controlled, and non-intrusive marking is used. In laboratories, mainly rats [35, 36], monkeys [37], and fish [38] are monitored. In some cases, insects [39, 40] and bats [41] are also monitored.

Livestock behavior monitoring is receiving more attention because it can be used to ensure animal welfare [8]. As in laboratory animal monitoring, the most

common aim is to identify, track, and sometimes categorize behavior. Pregnant cows' locomotion and posture were monitored in [42]. [43, 44] correlated low-level statistical measures to chicken welfare. Pig tracking and monitoring in pens have attracted extensive attention. Pigs in each image can be identified using markers [45], features [46] or applying shape matching technique [47]. Monitoring pigs using trackers is the most stable approach. [46] successfully tracked three pigs for eight minutes in a constrained area using features extraction method.

Multiple methods have been proposed for monitoring animal and human behaviors. Choosing the most optimal method depends on several factors, including the objective, monitored the environment, objects and their densities, and visibility.

Chapter 3

Multivariate Statistics

In this section, two multivariate statistical methods for behavior classification are discussed: *Principal Component Analysis* (PCA) and SVM. Both methods work well with correlated and high-dimensional data. PCA has already been used in monitoring processes and SVM has been used for a large variety of data classification problems, including in behavior analysis studies. In addition, the *Locally Scatterplot Smoothing* (LOESS) is presented, which is used for smoothing noisy observations.

3.1 Principal component analysis

PCA [48] is a statistical method that maps correlated data into another space where the data is uncorrelated. An orthogonal linear transformation is employed, and a new coordinate system is built such that the first axis in the transformed space is the linear projection of the greatest data variance and the second axis is the second greatest variance, etc. The new axes are named the *principal components*. The transformation of the data matrix X is defined as:

$$T = XW \tag{3.1}$$

where W is the *weights* or *loadings* and T is the *principal component scores*.The number of principal components is usually chosen using a threshold of the ac-

cumulated variance explained by the components.

PCA is commonly used as a primary tool in the investigatory analysis of data to highlight patterns and reduce dimensions. Sometimes more advanced methods are built based on PCA. In image analysis, PCA is used in face recognition [49, 50], feature extraction for crowd monitoring [51], and trajectory segmentation [52]. PCA is also applied to monitor the quality of manufacturing processes where multiple variables are monitored [53, 54].

In this thesis, PCA is used to monitor the movement of pigs in the *thermal data set*. The results of the analysis are presented in Paper A and Section 7.1.

3.2 Support Vector Machines

SVM [55, pp. 417-458] is a supervised classification method that maps data into a higher dimensional space where the classes can be separated using hyper-planes. SVM can be adapted for regression problems as well, but it will not be discussed further here because the problem analyzed in this thesis is classification.

Let a training data set be defined as $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$ where $x_i \in \mathbb{R}$ and $y_i \in -1, 1$. The hyper-plane separating the two classes is defined by:

$$\{\mathbf{x} : f(\mathbf{x}) = x^T\beta + \beta_0 = 0\} \tag{3.2}$$

where $\|\beta\| = 1$. The classification is performed by:

$$G(\mathbf{x}) = \text{sign}(\mathbf{x}^T\beta + \beta_0) \tag{3.3}$$

$G(\mathbf{x})$ is the signum function of the distance from $\mathbf{x}$ to the hyper-plane. In this thesis, $f(\mathbf{x})$ is referred as *Classification Decision Value* (CDV). If the classes are perfectly separable then $f(\mathbf{x})$ can be defined such that $y_i \cdot f(x_i) > 0 \quad \forall i$; therefore, the data is separated with the largest margin. The SVM is an optimization problem where the margin $\frac{1}{\|\beta\|}$ between the two classes is maximized. The optimization problem for non-overlapping classes is:

$$\begin{gathered} \arg\min_{\beta,\beta_0} \|\beta\| \\ \text{subject to } y_i(x_i^T\beta + \beta_0) \geq 1, i = 1, ..., N \end{gathered} \tag{3.4}$$

Normally the two classes cannot be perfectly sepereated and the above optimization will not be enough. Let ξ_i be a *slack* variable that defines the degree of misclassification of data point x_i:

$$y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i \tag{3.5}$$

The optimization problem is then a trade-off between a large margin and a small misclassification:

$$\arg\min_{\beta,\beta_0}\|\beta\| \text{ subject to } \begin{cases} y_i(x_i^T\beta+\beta_0) \geq 1-\xi_i, \forall i \\ \xi_i \geq 0, \sum \xi_i \leq C \end{cases} \tag{3.6}$$

where C is a constant.

When the data cannot be separated directly using a linear hyper-plane then more complicated separation functions must be used. The nonlinear classification is achieved by using kernel functions to map the data into another feature space where the linear hyper-plane will separate the classes. The following kernel functions have been used in the literature:

$$\begin{aligned} d^{th} \text{ Degree polynomial: } & K(\mathbf{x},\mathbf{x}') = (1+\langle \mathbf{x},\mathbf{x}'\rangle)^d \\ \text{Radial basis: } & K(\mathbf{x},\mathbf{x}') = \exp(-\gamma\|\mathbf{x}-\mathbf{x}'\|^2) \\ \text{Neural network: } & K(\mathbf{x},\mathbf{x}') = \tanh(\kappa_1\langle \mathbf{x},\mathbf{x}'\rangle + \kappa_2) \end{aligned} \tag{3.7}$$

To identify the best kernel function and parameters cross-validation is the most commonly used method. Some benefits of the SVM classification method include the following:

- Classes are separated with the largest margin.
- Linear and non-linear separation functions can be employed.
- The method can deal with correlated variables.
- The method is used for outlier and novelty detection.

One drawback of the SVM method is that it can only solve binary classification problems. This drawback is overcome by splitting multi-class problems into multiple binary problems. Another drawback of this method is that the structure of SVM cannot be interpreted.

SVM is applied in many areas. In computer vision, the method is used for human action recognition [56], face recognition [57, 58], gene analysis [59], and crowd density estimations [60]. In this thesis, SVM is used for pig behavior classification. The results are discussed in Papers B, C, D and Chapter 7.

3.3 Locally weighted scatterplot smoothing

Locally Weighted Scatterplot Smoothing (LOWESS) or LOESS ([55, p. 191-198], [61, p. 309 - 376]) is a non-parametric regression method used to smoothing scat-

terplots. A linear or non-linear least square regression model fit fitted locally to a subset of the data. The fitted functions over this local subset is a polynomial of degree up to 2. In addition, the weights can be chosen to include the data from the local subset with different proportions. The LOWESS solves the optimization problem at each point x_0:

$$\max_{\alpha(x_0),\beta_j(x_0),j=1,\ldots,d} \sum_{i=1}^{N} K_\lambda(x_0, x_i) \left[y_i - \beta_0(x_0) - \sum_{j=1}^{d} \beta_j(x_0) x_i^j \right]^2 \tag{3.8}$$

where $K_\lambda(x_0, x_i)$ is a weight function, $(x_0; y_0)$ data point from the local subset and $\beta_0(x_0) + \sum_{j=1}^{d} \beta_j(x_0) x_i^j$ polynomial of order d.

The use of LOESS gives advantage of not having a single model for the entire data. It is useful, when a single model approach does not provide smoothing due to local dynamics of in the data.

Chapter 4

Statistical process control

SPC is widely used to control and improve the quality of processes by variance reduction and specific cause elimination. This is usually achieved through a so-called control chart. The process monitoring is done in two phases: *phase I* refers to a chart estimation state using historical data, and *phase II* is an online monitoring state.

The main approach to cluster SPC control charts is based on the monitored characteristics' type and quantity as well as expected shifts in the process. Common characteristics in the manufacturing process, such as temperature, pressure, diameter, and weight, are usually expressed numerically on a continuous scale. However, some characteristics are only expressed in two classes: *conforming* or *non-conforming*. Such characteristics are known as *attributes* and the charts *atribute control charts*. In some cases, processes are complex, and multiple characteristics are monitored. *Univariate control charts* are used when one characteristic is monitored. If several characteristics are monitored, the charts are known as *multivariate control charts*.

A certain amount of variability is always present in all stable processes. If a process is operating only under natural variability, which is known as *variation due to chance causes*, then the process is said to be *in-control*. Variability that is not natural to the process is generated by *assignable causes*. A process is said to be in an *out-of-control* state when assignable causes are present in the process.

4.1 Chart establishment

Typically, a control chart consists of: (1) *Centet line*, (2) *Lower Control Limit* (LCL), and (3) *Upper Control Limit* (UCL). The center line represents an average in-control value of quality characteristics. The LCL and UCL are the bounds around the center line indicating acceptable levels of variability. A strong connection exists between control limits and statistical hypothesis testing. Choosing control limits corresponds to choosing a critical region in hypothesis testing. The general framework of a control chart is as follows:

$$\begin{aligned} UCL &= \mu_w + L_{cl}\sigma_w \\ \text{Center line} &= \mu_w \\ LCL &= \mu_w - L_{cl}\sigma_w \end{aligned} \tag{4.1}$$

where μ_w is the mean value of a quality characteristic, σ_w is the standard deviation of a characteristic, and L_{cl} is the distance from the center line to the control limits.

As mentioned in the introduction to this chapter, there are two phases in for process monitoring. In *phase I*, historical data that is believed to be in control is collected. It is assumed that data is in-of-control in *phase I* to ensure that only in-control data is included in chart estimation for *phase II*. Any out-of-control points detected in *phase I* control chart are investigated and removed if necessary. After the data is cleaned and additional data is collected if necessary, the final control limits are established for online monitoring.

Two types of errors are associated with determining whether processes are in-control or out-of-control. A *type I* error is made when a process is declared out-of-control, when it is actually in-control. A *type II* error is made when a process is declared in-control, when it is actually out-of-control. In some cases, the probability of a *type II* error is analyzed for different charts to indicate the size of the shift that the chart can detect.

The charts performance are commonly compared using *Average run-length* (ARL). The ARL_0 for in-control state is defined:

$$ARL_0 = \frac{1}{p} \tag{4.2}$$

where p is the *type I error*. If the *type I error* is fixed during in *phase I*, then ARL_0 of two charts can be compared during *phase II*. The average run-length for out-of-control state ARL_1 is defined as:

$$ARL_1 = \frac{1}{1-\psi} \tag{4.3}$$

where ψ is a probability of not detecting a shift on the first subsequent sample. For a good control chart it is expected, that the in-control ARL_0 is as large as possible while out-of-control ARL_1 as small as possible.

The principles of SPC are used to detect abnormal behaviors and locate abnormally moving pigs in a frame. The results are published in Papers A, B, D and chapter 7.

4.2 Univariate control chart

In most cases, the available data for monitored processes is continuous, and only one characteristic is available. A large variety of charts has been proposed for monitoring continuous processes. The majority of charts are built based on normality and independence assumptions. Appropriate charts are chosen based on the sample size and expected shift size. An overview of the charts is presented in Figure 4.1.

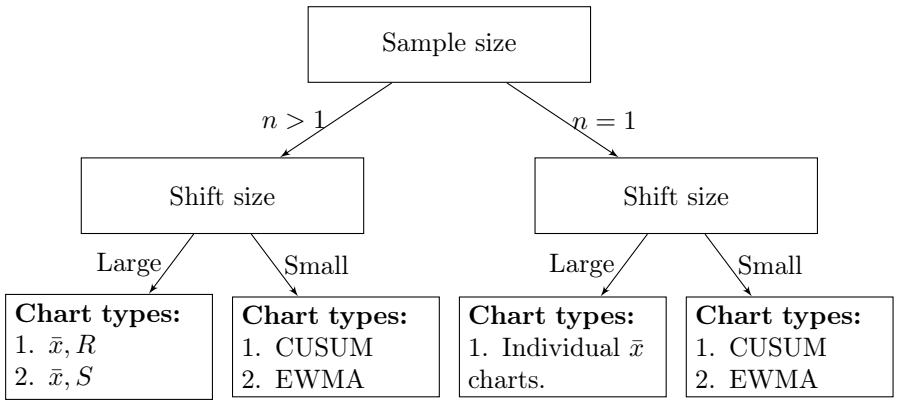


Figure 4.1: Flowchart for selecting the appropriate control chart for continuous characteristic monitoring.

Often it is not possible to inspect every produced object because the tests are either too time-consuming, expensive or destructive. In such cases, sampling is employed. The most commonly used chart for monitoring a sample mean is the $\bar{x}$ chart. $\bar{x}$ is used for smaller sample sizes, and it measures the average range of a sample. The $\bar{x}$ control chart for *phase I* is:

$$\begin{aligned} UCL &= \bar{\bar{x}} + A_2\bar{R} \\ \text{Center line} &= \bar{\bar{x}} \\ LCL &= \bar{\bar{x}} + A_2\bar{R} \end{aligned} \tag{4.4}$$

where the $\bar{\bar{x}}$ is the mean of the sample means; $\bar{R}$ is the mean of the sample ranges; and A_2 is a factor for control limits (the values can be found in [62, p. 700]). In most cases, $\bar{x}$ is used in conjunction with the $\bar{R}$ or $\bar{s}$ charts to obtain a better overview of the process. The $\bar{R}$ chart is:

$$\begin{aligned} UCL &= D_4\bar{R} \\ \text{Center line} &= \bar{R} \\ LCL &= D_3\bar{R} \end{aligned} \tag{4.5}$$

where D_3 and D_4 are factors for control limits. The $\bar{s}$ chart is:

$$\begin{aligned} UCL &= B_4\bar{s} \\ \text{Center line} &= \bar{s} \\ LCL &= B_3\bar{s} \end{aligned} \tag{4.6}$$

where B_3 and B_4 are factors for control limits and $\bar{s}$ is the average of sample standard deviations. Special adaptations of the chart are available to monitor varying sample sizes.

In some cases, the sample size is $n = 1$; therefore, individual measurements are available. In such cases, Shewhart control charts are applied, which are defined in Equation 4.1. If data is normally distributed, the moving range of two successive observations is used to estimate the process variability. The Shewhart chart for individuals is sensitive to violations of normality assumption. If normality is violated, the control limits should be derived according to the distribution or charts such as *Exponential Weighted Moving Average* (EWMA), which is less sensitive to normality assumption violations. The EWMA control chart, which are defined in Equation 4.7, is useful to detect small shifts and can handle auto-correlation.

$$\begin{aligned} UCL &= \mu_0 + L_{cl}\sigma\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})} \\ \text{Center line} &= \mu_0 \\ LCL &= \mu_0 - L_{cl}\sigma\sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2i})} \end{aligned} \tag{4.7}$$

where L_{cl} is the factor of control limit width, $0 < \lambda \leq 1$ is the smoothing parameter, and μ_0 is equal to the initial EWMA value $i = 0$:

$$EWMA_i = \lambda z_i + (1-\lambda)EWMA_{i-1} \tag{4.8}$$

The last chart mentioned in Figure 4.1 is the *Cumulative Sum* (CUSUM) chart. Several versions of the CUSUM chart are available, but the general framework

is:

$$\begin{aligned} C_{i-1}^{+} &= max[0, z_i - (M + K) + C_{i-1}^{+}] \\ C_{i-1}^{-} &= max[0, (M + K) - z_i + C_{i-1}^{-}] \end{aligned} \tag{4.9}$$

where C_{i-1}^{+} and C_{i-1}^{-} are the plotted statistics, M is the target value of z_i, and K is called "slack". When the values C_{i-1}^{+} and C_{i-1}^{-} exceed the predefined threshold H, the process is declared out-of-control. This chart is very useful for including the past history of the process in the decision making. The CUSUM chart accumulates deviations from the mean, and the chart signals when the accumulated values reach the predefined value.

The CUSUM chart is used to monitor pig behavior; the results are published in Paper D and chapter 7.

4.3 Attribute monitoring

Several charts can be used for attribute monitoring. The choice of the chart is based on two criteria: (1) whether the fraction of nonconforming units is analyzed or the count of nonconforming units and (2) the size of the expected shift. An overview of the attribute charts is presented in Figure 4.2.

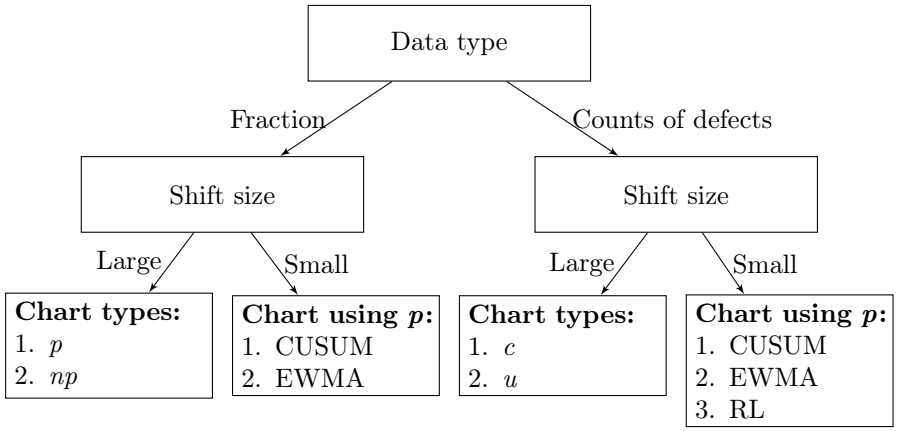


Figure 4.2: Flowchart for selecting the appropriate control chart for attributive characteristic monitoring.

The most basic attribute chart is known as the *p-chart* which measures the fraction of nonconforming units in a sample. The underlying distribution of this chart is Binomial. If the units in a sample are independent and n units are

included in the sample, D nonconforming units are expected in the sample. D follows a Binomial distribution. The sample nonconforming fraction is estimated as follows:

$$\hat{p} = \frac{D}{n} \tag{4.10}$$

To establish the chart, m good process samples must be collected, where m is at least $20 \leq m \leq 25$. Three standard deviation control limits for the *p-chart* are:

$$\begin{aligned} UCL &= \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \\ \text{Center line} = \bar{p} &= \frac{\sum_{i=1}^{m} D_i}{mn} \\ LCL &= \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \end{aligned} \tag{4.11}$$

The sample size must be considered carefully because it can affect the false alarm rate. Usually, large sample sizes are required for small $\bar{p}$ values. One technique used to choose the sample size is based on the desired shift size Δ that should be detected quickly:

$$n \geq \begin{pmatrix} L_{cl} \\ \Delta \end{pmatrix}^2 \bar{p}(1-\bar{p}) \tag{4.12}$$

The *p-chart* can be modified to accommodate the varying sample size.

In some cases, the actual number of nonconforming units in a sample is easier to interpret than the fraction. To monitor the count of nonconforming units, the *np-chart* is employed. If the size of the sample n is constant, the control chart is defined as:

$$\begin{aligned} UCL &= n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})} \\ \text{Center line} &= n\bar{p} \\ LCL &= n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})} \end{aligned} \tag{4.13}$$

More than one defect is often possible in a unit. A unit is declared as nonconforming if the unit has at least one defect (nonconformity). A chart showing the total number of nonconformities in a unit or the average number of nonconformities per unit can be applied. These charts are known as *c-chart* and *u-chart*, respectively. These charts are based on the Poisson distribution:

$$p(z) = \frac{e^{-c} z^c}{!z} \tag{4.14}$$

where z is the number of nonconformities. The mean and variance of the Poisson distribution is a parameter $c > 0$, so the *c-chart* control chart for the total

number of nonconformities is:

$$\begin{aligned} UCL &= \bar{c} + 2\sqrt{\bar{c}} \\ \text{Center line} = \bar{c} &= \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}}{m} \\ LCL &= \bar{c} - 2\sqrt{\bar{c}} \end{aligned} \tag{4.15}$$

The *u-chart* is used when the number of inspected units per sample varies.

$$\begin{aligned} UCL &= \bar{u} + 2\sqrt{\bar{u}} \\ \text{Center line} = \bar{u} &= \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}}{mn} \\ LCL &= \bar{u} - 2\sqrt{\bar{u}} \end{aligned} \tag{4.16}$$

As shown in Figure 4.2, the charts described above are most suitable for detecting shifts larger than 1.5σ. Extensions exists to the *p-chart*, *c-chart*, and *u-chart* to detect smaller shifts than 1.5σ in a process. These extensions are based on the EWMA and CUSUM control charts presented in Section 4.2.

An alternative approach to monitor an attribute process is to use *Run-Length* (RL) between two nonconforming units. [63] presents a RL-CUSUM chart when the time between events follows an exponential distribution. [64] presents RL-CUSUM and RL-Shewhart charts to find the sum of two recent RLs.

All the methods described above assume that the samples are independent. In some cases, this assumption is violated, and using standard methods increases the false alarm rate. Several attempts have been made to develop control charts for correlated attributes. [65] derives more suitable control limits for the RL chart for the sum of two recent RLs charts. [66, 67] suggest using a two-state Markov chain to construct an attribute control chart.

The attribute control charts presented in this section are not employed in this thesis. It is, however, necessary to be familiar with the principles of attribute control charts to understand the choices of some methods in this thesis.

4.4 Auto-correlated quality characteristics

The common assumptions for control charts are that the process follows normal distribution $N(\mu, \sigma)$ and samples are independent. Some control charts can tolerate violations of normality to a certain degree but not the violation of the independence assumption. The violation of independence usually results in increased numbers of false alarms. Two main approaches are available to deal with an auto-correlated process: (1) model-based and (2) model-free.

The time-series models are the most common way to deal with serial correlations in a process. A process can be modeled using a *Autoregressive Moving Average* (ARMA) model. The ARMA(p, q) of the auto-regressive order p and moving average order q is:

$$z_t = \varepsilon_t + \sum_{i=1}^{p} \phi_i z_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{4.17}$$

where ε_t is white noise, ϕ_i are the coefficients of the auto-regressive part, and θ_j are the coefficients of the moving average part. The ARMA model is validated by analyzing the residuals. Equation 4.17 is used to predict the process characteristic z at time t. Denote $\hat{z}_t$ as the one-step prediction of a characteristic at time t. The error between the predicted and observed characteristics at time t is:

$$e_t = z_t - \hat{z}_t \tag{4.18}$$

The one-step prediction error is monitored using the standard univariate control charts described in Section 4.2. The residuals should be normal and independently distributed with mean zero and constant variance. EWMA can also be used to monitor an auto-correlated process.

The model-free approach is based on un-weighted batch average control charts. Auto-correlated observations are divided into batches, and the average of each batch is used for monitoring. The most important step in this approach is determining the batch size. In general, the batch size must be large enough to decrease the auto-correlation between the batch averages to acceptable levels. The batch averages are monitored using the univariate control charts described in Section 4.2.

The data analyzed in this thesis is highly auto-correlated. The results of the monitored auto-correlated data are presented in Paper D and chapter 7.

4.5 Multivariate control charts

Some processes are complicated and defined by multiple characteristics. Monitoring each characteristic using individual charts is not optimal because it demands a lot of attention. In addition, in some cases, only combined characteristics indicate an out-of-control process, while individual characteristics do not due to potential correlation among these characteristics. To deal with such situations, a group of charts known as *multivariate control charts* [68] have been developed. The classical approaches (*Hotelling* T^2 *control chart* and χ^2 *control*

chart) defined in equations 4.19 and 4.20, respectively, are commonly used.

$$\begin{aligned} T^2 &= n(\bar{\mathbf{x}} - \bar{\bar{\mathbf{x}}})' S^{-1} (\bar{\mathbf{x}} - \bar{\bar{\mathbf{x}}}) \\ UCL &= \frac{w(m-1)(n-1)}{mn - m - w + 1} F_{\alpha, w, mn-m-w+1} \end{aligned} \tag{4.19}$$

where S is the estimated covariance matrix, $\bar{\bar{\mathbf{x}}}$ is a vector of in-control characteristic average values, n is the sample size, w is the number of quality characteristics, and m is the number of subgroups.

$$\begin{aligned} \chi^2 &= n(\bar{\mathbf{x}} - \mu)' \sum^{-1} (\bar{\mathbf{x}} - \mu) \\ UCL &= \chi^2_{\alpha, w} \end{aligned} \tag{4.20}$$

where $\sum$ is the known covariance matrix and μ is a vector of the mean values of the in-control characteristics. The LCL for both charts is equal to 0. The difference between these charts is that the χ^2 *control chart* is used when the covariance matrix and means are known for a process, and the *Hotelling* T^2 *control chart* is used when they are estimated from a process. For this reason, the *Hotelling* T^2 *control chart* has different control limits for *phase II* than those presented in Equation 4.19. In summary, these charts combine multiple characteristics to declare an out-of-control state. When the process is declared out-of-control, individual charts can be analyzed to detect the possible cause.

Other approaches are e.g. a multivariate EWMA control chart, regression adjustment, and multiway PCA for batch monitoring [53, 69, 70]. Only the latter will be presented because it is the only one used in this thesis. The multivariate batch process data is three-dimensional; J characteristics are observed over time K and L batches are recorded. JxK data is collected for each batch, resulting in a 3D data array $X(LxJxK)$. [53] suggests unfolding this data into 2D and apply an ordinary PCA. This unfolding is performed such that vertical slices LxJ are stacked next to each other, resulting in a transformed data matrix $X_{trans}(LxJK)$. The unfolding process is presented in Figure 4.3. The PCA is

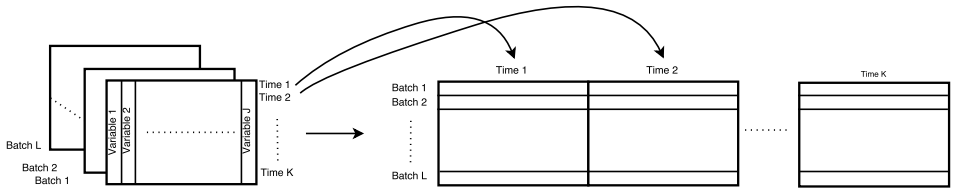


Figure 4.3: Unfolding data matrix

used on the transformed data. Typically only a few principle components will be needed to explain the majority of the process variation. Three charts are monitored to fully understand the process:

1 a sum of the square of residuals (SSE) of individual batches

$$SSE_i = \sum_{k=1}^{K} \sum_{j=1}^{J} \varepsilon(i, kj)^2 \tag{4.21}$$

2 elliptical contours in the joint chart of the first two principal components. The center coordinates of the contours are $(0, 0)$ and their axis length in the direction of r^{th} principal component in *phase I* are:

$$\pm S(r, r) B_{1, \frac{L-2-1}{2}, \alpha} \sqrt{\frac{(L-1)^2}{L}} \tag{4.22}$$

3 the *Hotelling* statistics T^2:

$$T^2 = t'_Q S^{-1} t_Q \frac{L}{(L-1)^2} \sim B_{\frac{R}{2}, \frac{L-R-1}{2}, \alpha} \tag{4.23}$$

where t_k is the score of the k^{th} principle component, Q is the number of principal components used for the analysis, L is the number of historical in-control batches, R retained number of pricipal components, and B is beta distribution.

The analyzed process in this thesis is multivariate. PCA-based control charts are used in Paper A. CUSUM chart is used for motingoring SVM CDV and results are presented in Paper D and chapter 7.

Chapter 5

Markov chains

A *Markov Chains* (MC) are a sequence of random variables which transitions from one state to another in a finite state space. The transition to the next state only depends on the current state. Only MC is introduced here and an extended discussion of both MC can be found in [71].

A sequence of discrete random variables $\{z_t : t \in \mathbb{N}\}$ is a MC if it satisfies property:

$$Pr(z_{t+1}|z_t, ..., z_1) = Pr(z_{t+1}|z_t) \tag{5.1}$$

Equation 5.1 means the next random variable is conditioned only on the most recent variable. The most important characteristic of MC is the *transition probabilities*:

$$\vartheta_{ij}(t) = Pr(z_{s+t} = j|z_s = i) \tag{5.2}$$

Equation 5.2 is called the homogeneous probabilities, which means the transition probability does not depend on time. $\Gamma(t)$ is a *transition matrix* $\vartheta_{ij}(t)$ is the probability of a process to move from state i to j. The transition matrix has the following properties:

- The row of transition probabilities sums to 1.

- The r^{th} transition of the transition probabilities is defined using the **Chapman-Kolomogorov equation** for conditional probabilities:

$$\Gamma(t+r) = \Gamma(t)\Gamma(r) \tag{5.3}$$

- The transition of unconditional probabilities is:

$$r(t+1) = r(t)\Gamma \tag{5.4}$$

- The Markov chain is stationary if it has a stationary distribution τ, which satisfies

$$\begin{aligned} \tau\Gamma &= \tau \\ \tau 1' &= 1 \end{aligned} \tag{5.5}$$

- The MC is said to be reversible if:

$$\tau_i \vartheta_{ij} = \tau_j \vartheta_{ji} \tag{5.6}$$

- If it is possible to reach all states from all states in the process, then Γ is said to be irreducible.

- If a m state MC $\{z_t\}$ is stationary and irreducible, the auto-correlation function for lag k is:

$$\rho(k) = Corr(z_t, z_{t+k}) = \frac{\sum_{i=2}^{m} g_i h_i w_i^k}{\sum_{i=2}^{m} g_i h_i} \tag{5.7}$$

  where w_i are eigenvalues and can be estimated using the eigendecomposition.

The transition probability matrix is estimated element by element:

$$\vartheta_{i,j} = \frac{f_{i,j}}{\sum f_{i,\cdot}} \tag{5.8}$$

where $f_{i,j}$ is the number of transitions from state i to state j and $i, j = 1, ..., m$. It can be shown that Equation 5.8 is the maximum likelihood estimation of Γ conditioned on the first observation.

Markov chains have previously been used for auto-correlated attribute monitoring [72, 66]. The Markov chain is also used in behavior classification problems [32, 32]. I propose to use MC for assignable cause identification. The results are presented in Sections 6.7 and 7.8.

Chapter 6

The use of image analysis for statistical process control

This chapter presents the proposed framework for ensuring the welfare of pigs at a slaughterhouse using video surveillance. The overall real-time monitoring scheme is presented in Section 6.1. The individual steps and the method training are discussed in the following sections of this chapter.

6.1 Methodology scheme

It is common to use video surveillance to ensure public safety and well-being. The general framework for surveillance and object tracking is presented in section 2.4. Pigs are commonly tracked using either markers, extracted features or shape matching, and are usually in a constrained environment, e.g. pens. The tracking approach is not suitable for the present thesis. Tracking a large number of pigs, which are able to leave and re-enter the surveyed area using extracted features can be computationally expensive. Individually marking pigs is the most efficient method for tracking, yet this is not possible in slaughterhouses

due to the large number of pigs. The shape-matching method can become ineffective when dealing with large numbers of pigs as the outlines disappear if pigs are in close proximity. [46] first proposed the method for monitoring pigs without using markers or features, though pigs are in a constrained environment. Monitoring individual humans in large dynamic crowds without special markers or trackers is rarely attempted when subjects can leave and re-enter the surveyed area. The OF is the preferred approach for monitoring crowd movement in such cases. Large crowds are common at public events, where humans move in an unconstrained environment. Most commonly, the entire crowd movement is analyzed to detect abnormal behavior without identification of individuals. In this thesis, the pig herd is monitored using the human crowd approach, as individual pigs are difficult to identify due to their similarities and density. In general, human crowd monitoring involves the following steps: motion estimation, object identification, feature extraction, event modeling and classification.

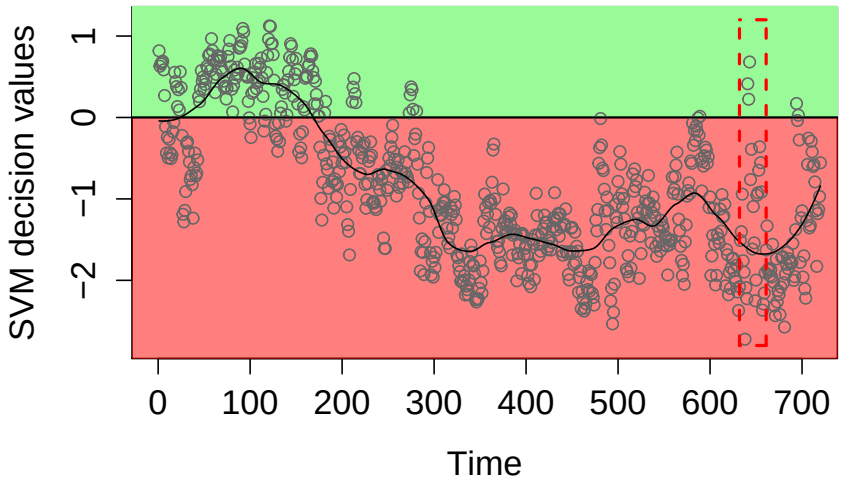


Figure 6.1: The illustration of discrete and continuous SVM. The grey points are CDV, the black line is smoothed CVD, the red dash rectangle indicates an event, the green background indicates in-control events based on discrete SVM and the red background indicates out-of-control events.

A detailed description of OF, which is used for motion estimation, is presented in section 2.1. There are multiple methods suggested in the literature for object identification in a frame. The literature review of object identification is presented in section 2.2. The choice of method is usually dependent upon the visual properties of the surveyed objects, background illumination and other factors. There are multiple feature extraction methods proposed in the literature, and the overview is presented in section 2.4. There are vast differences between the body shape and movement of pigs and humans. In this thesis, the *Modified Angular Histogram* (MAH) was used to summarize the pig herd movement. The difference between pig and human movement is discussed in section 6.3. Event modeling and classification is performed in the final surveillance step. In some cases, features are directly used for classification. An overview of possible methods for human crowd modeling and classification is presented in section 2.4. In this thesis, the method used for classification of the features is SVM (see section 3.2), which is commonly used for dense crowd monitoring and can handle correlation between the variables.

Motion is continuous, where one event merges into another event over time. In most cases, the continuation of the motion is disregarded when using standard computer vision methods for surveillance. In this thesis, the continuous classification values together with the principles of SPC are used in decision making to accommodate the continuity of motion. The majority of classification methods make a final decision based on probabilities, scores or other continuous values. Figure 6.1 illustrates the SVM decision-making process. The grey points are SVM CDV that gradually drift from one state to another and are highly auto-correlated. In a standard discrete approach, if an event is detected then CDV are above or below 0. Utilizing the continuity and auto-correlation of CDV can result in earlier detection of events and a decrease in misclassification. A smoothing method is applied to reduce noise that can result in misclassification. The principles of SPC are employed to make a final decision based on a CDV that allows early event detection and a further decrease in misclassification. The performance comparison using discrete and continuous SVM is presented in section 7.7. The proposed real-time monitoring framework is presented in Figure 6.2.
In an analyzed case, declaring an event is not sufficient; it is also important to locate the event in a large area. In this thesis, the location of the undesirable event in a surveyed area based on principles of SPC and MAH is presented. The detailed procedure is presented in section 6.6.

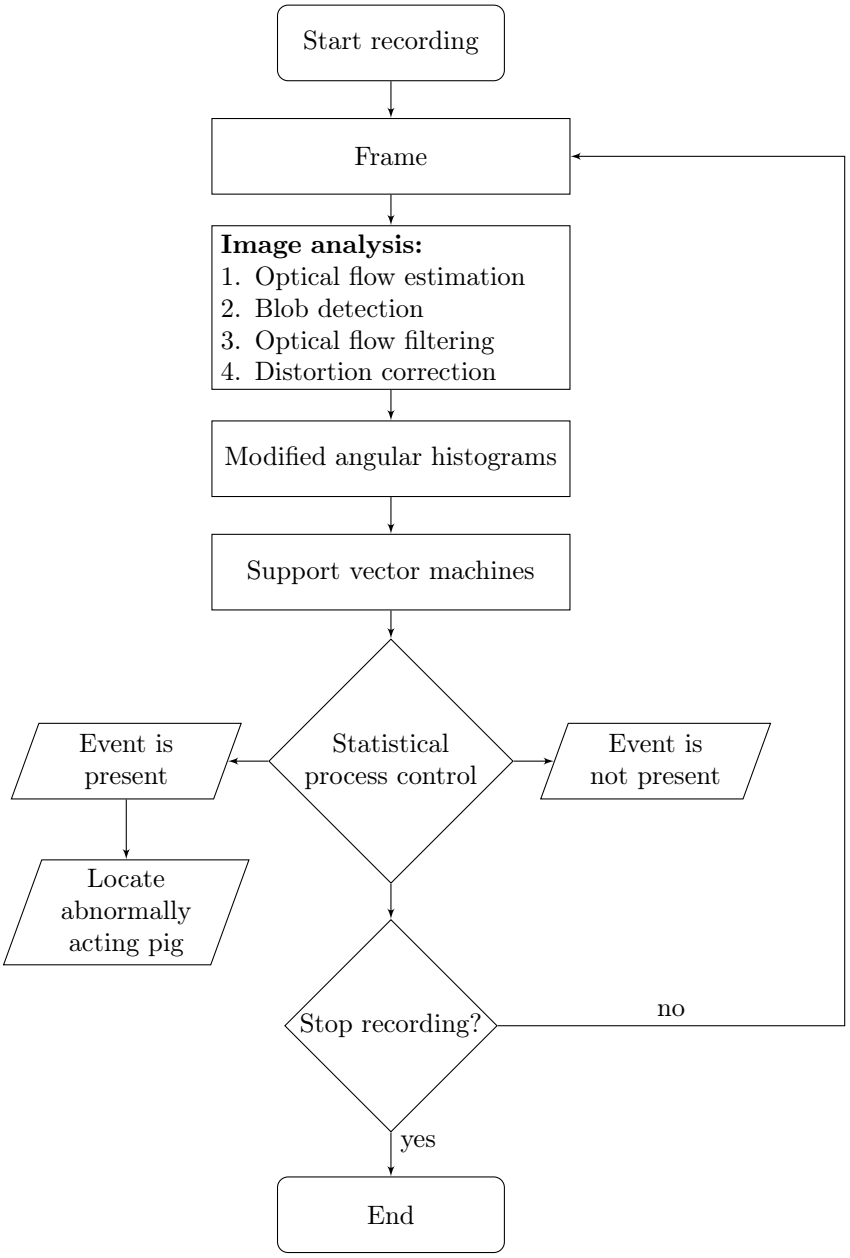


Figure 6.2: Flowchart for monitoring the behavior of pigs in a slaughterhouse.

6.2 Methodology for lens and foreshortening distortions correction

The first two steps in surveillance are motion estimation and object detection. In this thesis, OF is employed to capture the motion of the pigs, and the background subtraction, together with blob detection are used to identify pigs in a frame. An additional step must be added to the procedure, as two distortions: radial lens and foreshortening, are present in the videos. This thesis contributes a proposed procedure for the simultaneous correction of both distortions. In most published surveillance cases, only radial lens distortion is present. In such cases, the image is first corrected and then the motion is estimated. The proposed approach corrects the estimated OF field instead of correcting a frame. This method requires a large number of moving objects of approximately the same size to be recorded in a video. The measurements of these objects are used to build a correctional model. In this thesis, the pigs are all approximately 6 months of age and weigh 100-110kg.

In most cases, the magnitude of distortion depends on the position of an object in the surveillance area relative to a camera. For example, in the case of foreshortening distortion, objects near the camera appear larger. In radial lens distortion, objects that are closer to the focal point of a camera appear larger. The correction method proposed in this thesis is based on the assumption that OF vectors and the size of the objects are distorted proportionately to their distance to the focal point of the camera. This assumption is illustrated in Figure 6.3 for foreshortening distortion. It is assumed that objects are moving at the same speed in the illustration. The smaller the object is, the longer the OF vectors. Using this relation, the OF vectors are corrected instead of the image. The correction is achieved by recording the length and width of the objects, so the measurements are perpendicular to the X and Y coordinate system. The measurements are assigned to the bottommost and leftmost points of an object, respectively. Multiple measurements are taken throughout the entire frame. The length and width measurements are modeled using a third-order polynomial which has also been used in [19] for distortion correction:

$$Z(x,y) = \sum_{j=1}^{4}\sum_{i=1}^{4} \beta_{ji} x^{4-j} y^{4-i} \tag{6.1}$$

The results of estimated length and width measurements for entire area of intres in *unloading data set* is presented in Figure 6.4. The defined field is used to correct lens and foreshortening distortions. The length and width $I_{lw} = \{m_l; m_w\}$ are estimated for every point in the image, before being rescaled to $[o; r]$ using

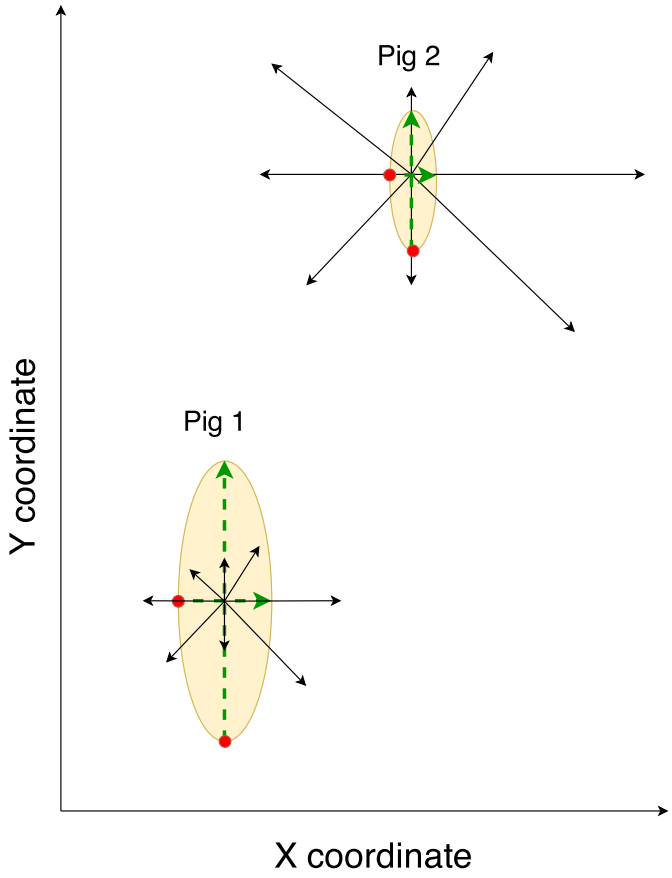
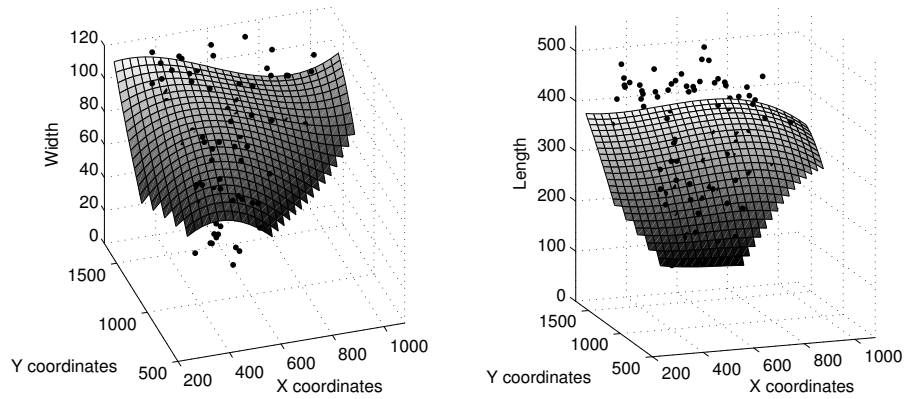


Figure 6.3: Illustration of the relation between OF vector lengths and pig sizes. Black arrows represent OF vectors, green arrows represent size measurements, and the measurements are assigned to the red points.

the following formula:

$$
\begin{aligned}
f(z) &= \frac{(o-r)(z-min(I_{lw}))}{max(I_{lw})-min(I_{lw})} + r \\
r &= 1 \\
o &= \frac{max(I_{lw})}{min(I_{lw})}
\end{aligned}
\tag{6.2}
$$



(a) Pig's length. **(b)** Pig's width.

Figure 6.4: The length and width of the pigs over the area of interest (marked as dots), and the 3^{rd} order model fit used for correction of lens and foreshortening distortions.

The distortion correction field is used to adjust the OF vectors. This is achieved by multiplying the real values of the OF vector with the width distortion correction coefficients and the imaginary value with the length distortion correction coefficients.

This simple approach can be applied to any type of distortion or combination of distortions. It is easy to implement the method, though it can only be used when references are available. The results and validation of the method are presented in Section 7.3.

6.3 Modified angular histograms

There are many human crowd movement feature extraction methods proposed in the literature (see section 2.4). None of these address the specific movement pattern of a single object in a group of large objects recorded in close proximity. The data resulting from such a pattern often contain noise, and identifying the movement direction and/or speed can be challenging. For example, humans do not have a specific walking pattern. As shown in [73], a polar histogram of crowd OF vectors has a single peak (with a moderate amount of noise) that represents the movement direction. As a result, current methods for monitoring human crowds usually assume that the OF vector angle values are Gaussian distributed.

Recent dense crowd-monitoring methods [74, 75, 76] use the principles of hydrodynamics. Each frame is represented by small particles, and the density and velocity of the particles are monitored in order to cluster and track movement. The speed and direction of a particle are the spatial-temporal averages of OF vectors over the surface area of the particle. An advantage of such methods is that they average out the noise created by the OF estimation. The principles of hydrodynamics give valuable results in the above papers.

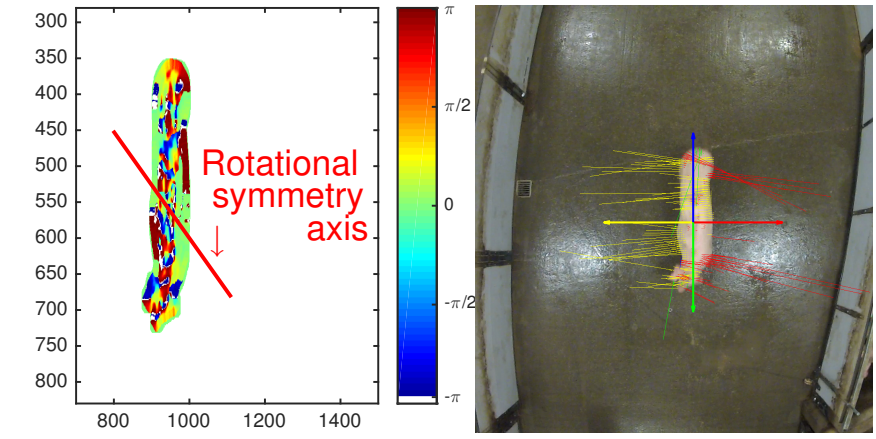
In this thesis, MAH was used to summarize the OF of the moving pig herd, and for motion analysis and event detection. The MAH is based on a polar histogram [73], built using one-degree intervals. To decrease the number of bins in the histogram, the Doane's formula was used to estimate the number of bins in the MAH:

$$k = 1 + \log_2(n) + \log_2\left(\frac{|g_1|}{\sigma_{g_1}}\right) \tag{6.3}$$

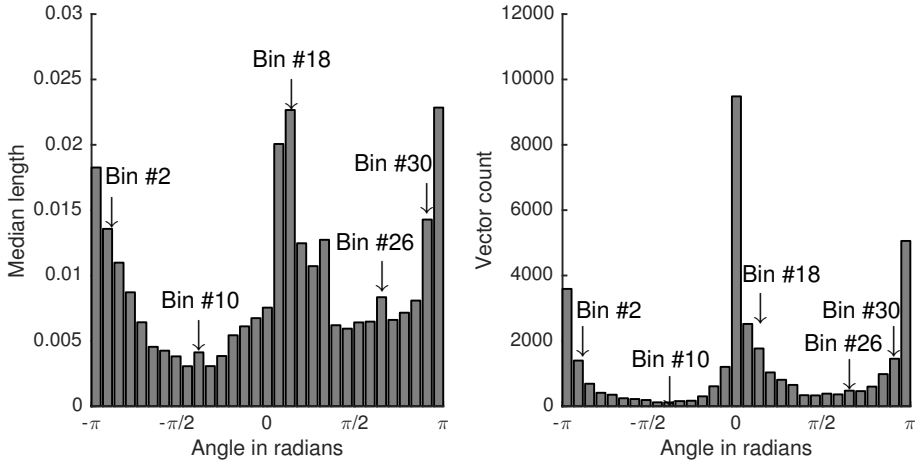
$$\sigma_{g_1} = sqrt(\frac{6(n-2)}{(n+1)(n+3)} \tag{6.4}$$

where g_1 is the sample skewness, n is the number of observations, and k is the number of bins.

Figures 6.5c and 6.5d indicate that the MAH of a pig herd's movement is bimodal. Pigs trot as they move forward giving the observed "S-shape" movement when recorded from above. Due to the trot and the elongated body shape, the OF vectors point in multiple directions. Figure 6.5a demonstrates that the angles of the OF vectors are approximately symmetrical to the rotational axis, indicating a trot. Furthermore, in Figure 6.5b, the vectors representing movement directions are not the longest, which means that the movement direction is in a valley of the MAH and not in a peak, which is around bin 26 in Figures 6.5c and 6.5d. The vectors of the movement direction are 2.7 times shorter than the two peaks. In addition, if we apply particle advection on the recordings, the true information of a movement's direction and speed may be lost since only 8% of all estimated OF vectors are caused by the movement direction. We assume that the direction of movement is from the top to the bottom of the image in Figure 6.5b. This means that particle advection or similar methods to analyze all vectors would most likely average out the vectors representing the true movement direction. The detailed results for the performance of MAH can be found in paper Papers B and C.



(a) Pig's OF angles. **(b)** Pig's OF.



(c) MAH of OF vector lengths. **(d)** MAH of OF vector counts.

Figure 6.5: Illustration of pig's OF, which moves from the top to the bottom of the image. Figure 6.5a indicates that vectors are pointing in multiple directions, and Figure 6.5b indicates that vectors in the direction of the movement are not the longest. The movement is summarized in MAH in Figures 6.5c and 6.5c.

6.4 Decision making

In surveillance, final decisions are often based on discrete classification. This means that classification is performed in extremes: either an event is present or it is not. However, in most cases, motion is continuous, and therefore an event can transition from one to another. In this thesis, the continuity of motion is achieved by analyzing continuous values of the classification method. Most classification methods assign the class based on continuous probabilities, scores, or other types of decision values. The SVM makes decisions based on continuous decision values CDV, which measure the distance to the separating hyperplane (see section 3.2). The sample of CDV is presented in Figure 6.1. The values are highly auto-correlated and noisy, thus smoothing must be applied. In statistical process control, the primary choice for monitoring noisy, auto-correlated data is the EWMA chart. The analysis indicates that analyzed data are homoscedastic, and therefore it is not possible to identify the optimal smoothing parameter. In this thesis, LOESS is used to smooth the data locally and does not require a specific model. Monitoring is performed using the CUSUM control chart, which collaborates previous observations in decision-making. As such, the transition from one action to another can be captured early on. In addition, the CUSUM chart was used in monitoring crowd congestion in [77]. The CDV is expected to be positive for in-control processes and negative for out-of-control processes in the analyzed data. Therefore, only the lower-limit of the CUSUM chart is of interest. Three parameters must be identified to establish the chart: (1) smoothing span N; (2) $(M + K)$; and (3) threshold h. The parameters can be chosen to optimize: (1) the total classification rate; (2) the early detection rate; or (3) based on the SPC approach. The early detection rate refers to the number of frames classified as out-of-control before the event is annotated. The results of the parameter selection are presented in section 7.7.

6.5 Method training

There are two phases in establishing a control chart (see chapter 4). In *phase I*, historical in-control data are collected and checked for outliers. In *phase II*, the process is monitored online. Collecting historical in-control data from videos is not always a straightforward process. In most cases, motion is a continuous process, where the beginning and end of actions or events are not clear. For example, according to the Oxford English Dictionary, a stampede is defined as, "a sudden panicked rush of a number of horses, cattle, or other animals" [78]. This definition raises two questions relating to annotations: (1) the number of animals and (2) the speed at which they must move. In video analysis, it is com-

mon to use human annotators to answer the questions "how many" and "how fast." However, several studies [79, 80, 81] indicate that human annotation quite often depends on personal understanding. [79] compares spatial, temporal, and behavioral annotation for human actions of different annotators. The image annotators had an agreement rate of 89%-97% for temporal action annotation. The majority of disagreement occurs at the beginning and end of the action, which is around 25-50 frames. The annotators agree on 80% of the action labels. The error rates are dependent upon the quality of the videos, the number of labeled actions, and the number of tracked objects. In our experience, the annotation becomes complicated when large crowds are monitored. The analysis in this thesis only includes two easily distinguishable actions. However, detecting the beginning and end of the action is a much more difficult task.

In the proposed framework for pig behavior monitoring, the parameters for the control chart and classification model must be established. The selection of the historical in-control data can affect the performance of both steps. [58] compares different representation approaches, features, and classification methods for gender classification problems. The results show that the SVM method using global features performs very well in classification problems, yet the same conditions must be present in the training and testing sets.

The literature for control chart parameter estimation is reviewed in [82]. They conclude that it is very important to select a sufficiently large, true representation of in-control data in *phase I* because it can affect the performance of the chart in *phase II*. An inappropriated training set will lead to the incorrect establishment of the distribution, and therefore to inappropriate control chart parameters in *phase II*. This might result in undetected shifts in a process and/or an increased number of false alarms. In addition, [82] states that the size of the sample influences the parameter estimation. A larger data set for *phase I* must be collected for a multivariate chart than for a univariate case. Where auto-correlation is present, the data set must be even larger than for the time independent data.

After evaluating the published results regarding annotation precision and the sensitivity of the models, it is recommended in this thesis that only unambiguous frames be used for the training set. As it is very difficult to numerically define a stampede, other indicators are used. The result of a stampede is often an event such as animals tripping over or stepping on each other. Therefore, only frames showing the event are used for the out-of-control training class. It is also difficult to identify whether pigs are moving "normally" when the driver is encouraging them. Thus, frames showing no personnel and no events are used as in-control training frames. For SVM model estimation, both in-control and out-of-control frames should be used, while only in-control frames should be used for establishing the control chart

6.6 Identification of stationary pigs

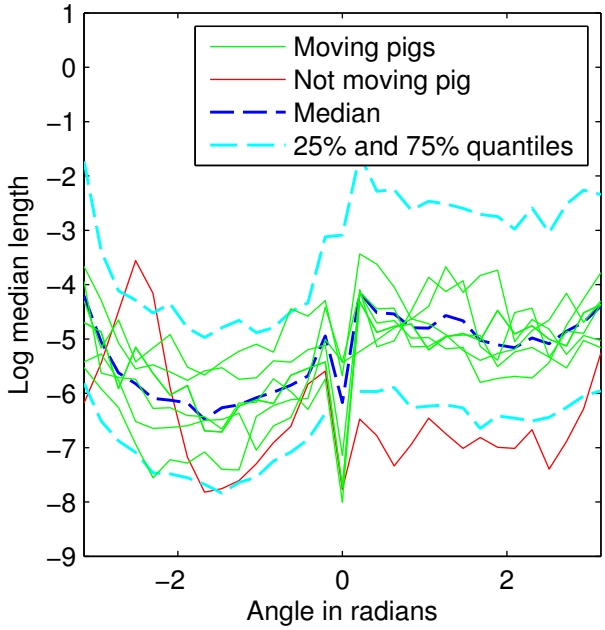


Figure 6.6: Identification of stationary pigs.

In this section, individual MAH of moving and stationary pigs are compared. In addition, the method for locating a stationary pig in a frame is presented. For illustration, all pigs are identified in one frame from an *entrance data set*. In this frame, all but one pig is moving. A MAH is built for each pig, and the median and 25th and 75th percentiles are estimated for all moving pigs. The comparison of MAHs is presented in Figure 6.6. The MAHs of moving pigs are within the 25th and 75th percentiles, while the MAH of the stationary pig is not. The comparison demonstrates that not all angles indicate whether a pig is moving. A sequential feature selection using SVM is used to identify the most important angles in the identification of a stationary pig. From *entrance data set*, a total of 540 pigs (294 stationary and 246 moving) are manually identified, annotated, and used for sequential feature selection. The results of sequential feature selection indicate that only four vectors are of great importance. All vectors point towards the direction of movement and are presented in Figure 6.7.

The procedure to identify a stationary pig in a frame is performed in two stages: (1) a training stage (which is offline) and (2) a testing stage, which can be implemented for real-time monitoring. In the training stage, the median and

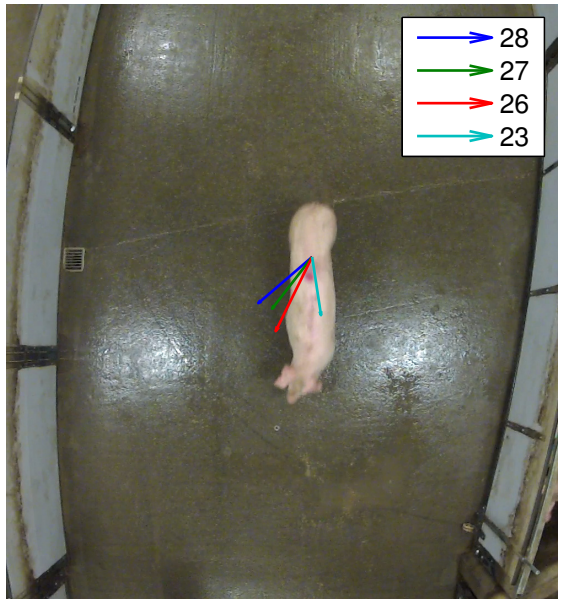
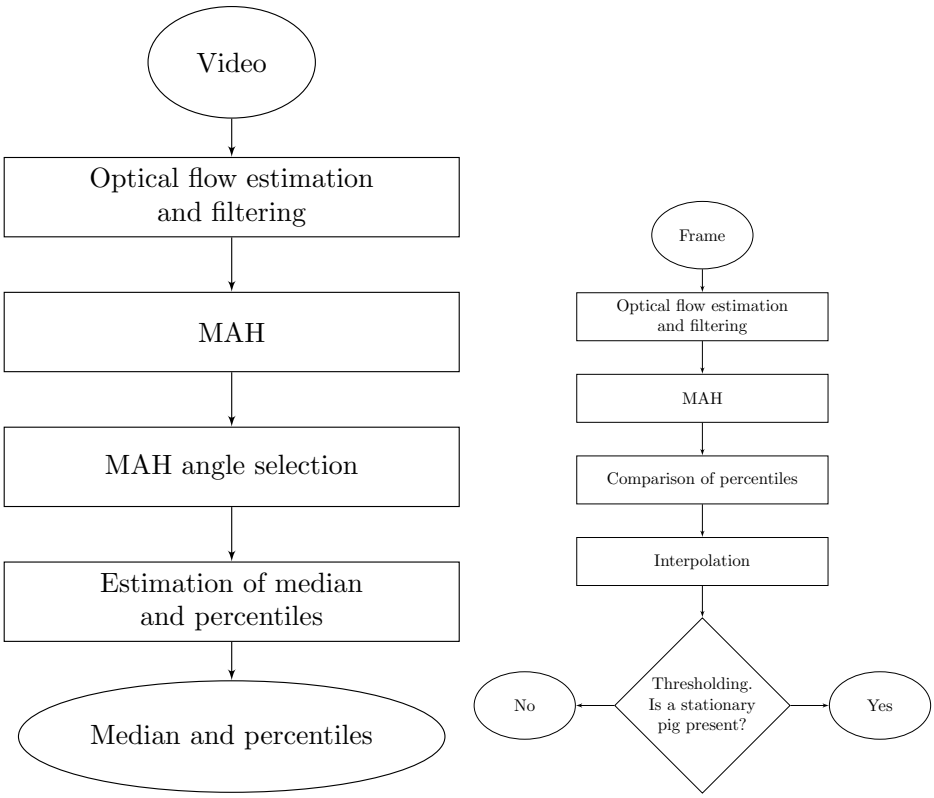


Figure 6.7: The results of sequential feature selection.

25th and 75th percentiles of selected angle bins of moving pigs are estimated. In the testing stage, only those OF vectors within selected angle bins are monitored. The vector is declared as out-of-control if it is not within the 25th to 75th percentile range. The density of the out-of-control vectors will be largest in an area where stationary pigs are present. The density is tracked over the patch of $61x61$ pixels. The threshold is employed to detect the out-of-control density. The overview of the training and testing stages is presented in Figure 6.8.

6.7 Assignable cause identification in auto-correlated process

In a classic SPC, binomial distribution is used to identify assignable causes. Binomial distribution requires independent observations, but those from recordings are highly auto-correlated. In this case, events can be defined using two parameters: frequency of occurrence and length of event. This thesis proposes a method based on MC to identify assignable causes.



(a) The training procedure. **(b)** The testing procedure.

Figure 6.8: Training and testing flowcharts of the method for stationary pig detection.

To define a two-state MC where *state I* is the in-control state, and *state II* is the out-of-control state, the transition matrix is as follows:

$$\begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \tag{6.5}$$

The probability that the process will stay in *state I* is p and in *state II* is q. Therefore, $(1-p)$ indicates how frequently the process moved from *state I* to *state II* or the frequency of the process becoming out-of-control. The q indicates the duration of the out-of-control state. If the transition matrices can be established under different circumstances, a comparison can be performed.

The methodology for estimating the transition matrix is well known and is widely used in many areas. The proposed method is promising, but additional

work must be done to validate it. The MC-based approach is compared to binomial distribution in Section 7.8.

Chapter 7

Results

7.1 Motion monitoring using PCA

The multivariate control chart based on PCA (see section 4.5) is used in the first instance to monitor pigs' motion in thermal data. Such a chart is established in two phases, as described in section 4.1. The objective of the analysis in *thermal data set* was to detect pigs that moved in the opposite direction, (i.e. from right to left) or that did not move at all. Each frame in a video was represented by two histograms: (1) a histogram of the pigs' OF vector angles and (2) a histogram of the pigs' OF vector lengths. OF was estimated for entire frame and then filtered using pixel color thresholding. The histograms were stacked next to each other and each frame was represented by a vector of length J. The information in the K frames was collected into a matrix, representing a scene. L scenes were collected and a 3D array obtained, representing a video. An illustration of the unfolding is presented in Figure 4.3. In *phase I*, an analysis of the sum of squared error, score plots, and Hotelling's statistics did not reveal any outliers. As suggested in the literature, only two *Principal Component*s (PCs) were considered for monitoring. In *phase II*, the new observations were collected and mapped in the PC space and the score plot was analyzed using equation 7.1.

$$\pm S(r,r)F_{2,L-2,\alpha}2\sqrt{\frac{L^2-1}{L(L-2)}} \tag{7.1}$$

where S is sample covariance matrix, r is index of principle component, F represents the F-distribution, and L is the number of historical in-control batches. A score plot of new observations is presented in Figure 7.1. The results indicate that most errors appear close to the control limits. The method correctly

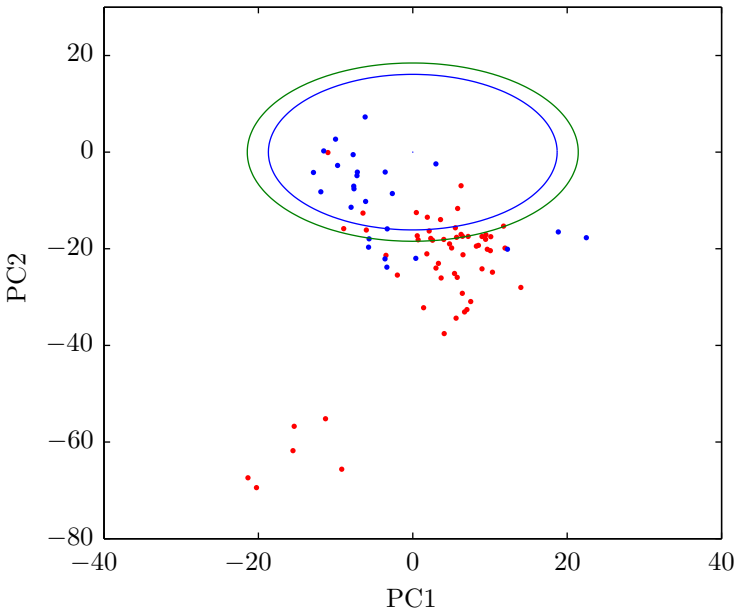


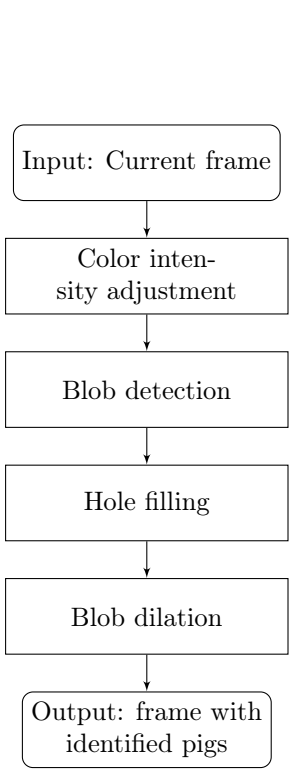
Figure 7.1: Results of Phase II. The blue dots represent normal movement, and red dots represent undesirable movement.

classifis 66% of the observations. There are several possible reasons for the moderately low classification rate:

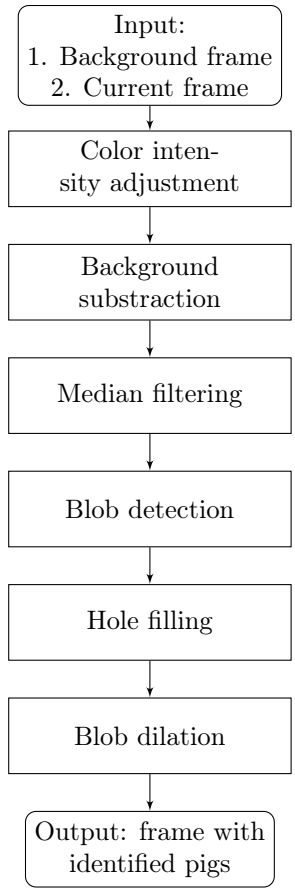
- Some errors could have accrued due to annotation. All the frames in the video were annotated. In some cases, it is difficult to determine where one action ends and another begins.
- Better classification accuracy can be achieved using additional principle components, but monitoring multiple charts is not optimal.

The multivariate PCA does not give the desired precision in pig motion monitoring. A different approach for frame annotation and other classification methods should be considered. Some additional results for this study case are presented in Paper A.

7.2 Pig identification in a color frame



(a) Blob analysis



(b) Background subtraction.

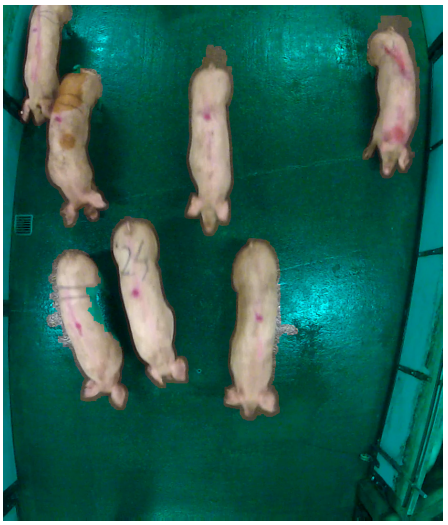
Figure 7.2: Flowcharts of blob analysis and background subtraction methods.

The first step in video surveillance is motion estimation and object detection as shown in flowchart 2.1. In this thesis, the motion of the pigs was estimated using the OF Horn-Schunck algorithm, implemented in Matlab [83] with a five-frame delay. Two approaches for pig identification in a frame were considered in this thesis. The approaches were based on: (1) blob detection and (2) background subtraction. The first approach applied blob detection directly to the original frame. The second approach first subtracted the background frame from the original and then identified blobs in the remaining frame. Some basic frame processing (image color intensity adjustment; hole filling; median filtering; and

blob dilation) was performed before and after subtracting the background frame, in order to improve the performance of both methods. An overview of the identification step is presented in Figure 7.2. The surveyed area in a frame was

(a) Unloading, background subtraction.

(b) Unloading, blob analysis.



(c) Entrance, background subtraction.

(d) Entrance, blob analysis.

Figure 7.3: Comparison of two methods for pig identification in a frame. The frames on the left represent the background subtraction method, while the frames on the right represent blob analysis. The top frames are from the pig unloading recordings, while the bottom frames are from the entrance recordings.

not affected by dynamic background or extreme illumination changes, though lighting reflections were visible in the *entrance* videos. The identification errors might have been caused by markings on the pigs' backs, skin coloring, or shadowing. It is expected that a combination of background subtraction and blob detection methods perform better in noisy frames.

The approaches were compared using two different types of frames from: *entrance data* and *unloading data sets*. The results are presented in Figure 7.3. Due to light reflection on the floor in the *entrance data set*, the background subtraction method performed better. The blob detection method classified light reflections as pigs. The blob detection method performed better for the *unloading data set*. The background subtraction method made some classification errors around the edges of the pigs' bodies. These results indicate that most efficient method for identifying pigs in the videos was dependent upon the surveyed area.

7.3 Lens correction

It is common in image studies in farms and slaughterhouses to use radial lens distortion, which is also used in the video recordings of the *unloading* and *entrance data sets*. The camera was positioned at an angle in the *entrance data sets*, thus creating additional distortion, which is called foreshortening distortion. As discussed in section 2.3, formulas have been proposed to correct individual distortions but not a combination of distortions. A method to correct a combination of distortions using moving reference measures and correcting the OF field instead of a frame is proposed in section 6.2. The direct validation of the method is not straightforward since the ground truth is unknown. Three indirect tests were used to validate the results:

1 Lens focal point estimation. The estimated focal point should match the visually observed focal point.

2 A single pig measurement. These measurements were recorded as the pig entered and left the area of interest. The measurements, such as length and width, were taken over 21 frames and corrected using the proposed model. The corrected measurements should be the same over 21 frames.

3 Classification performed with and without lens distortion. Three performance measures were compared: (1) total classification rate; (2) frames with event classification rates; and (3) frames without event classification rates. The data sets were randomly divided into training and testing

sets. The classification rates were computed for the testing sets and bootstrapped until the mean of each of the three rates converged, which occurred after 10 bootstrap samples. Analysis was performed using a *t-test* was. It was expected that lens correction would normalize the OF vectors over the entire frame, allowing events to be detected independently of the event location in a frame.

The results of the first indirect test are presented in Figure 7.4. The lens focal point is where the distortion was smallest, and is represented by the yellow circuit. Visual inspection also suggested that the lens focal point was in approximately the same location.

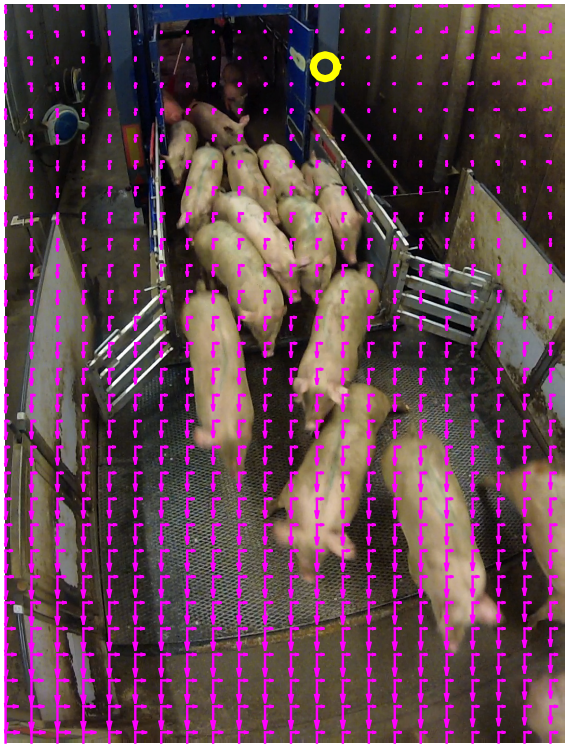
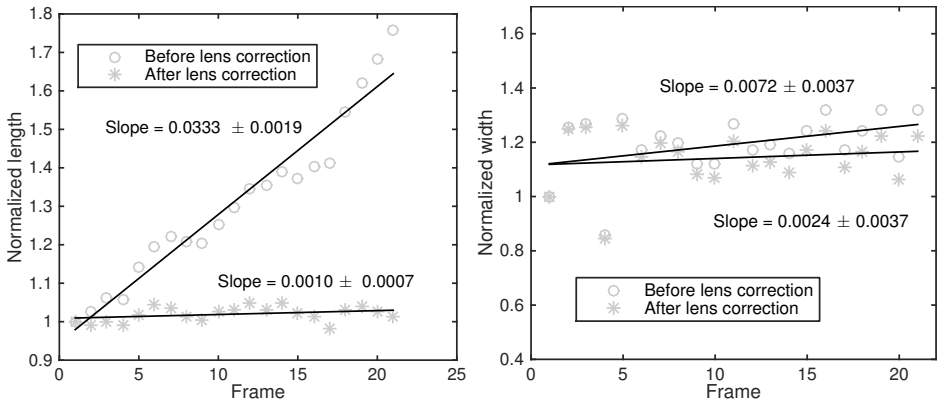


Figure 7.4: Detected lens focal point. Purple vectors represent the magnitude of the distortion, and the yellow circuit represents the estimated lens focal point.

The recorded and corrected measurements for a single pig are presented in Figure 7.5. The length and width measurements of the pig before correction increased approximately linearly, and after correction, they were assumed to be constant.



(a) Pig's length. **(b)** Pig's width.

Figure 7.5: The length and width of one pig sampled over 21 frames. In frame 1, the pig is fully visible in the area of interest for the first time; in frame 21, it is visible for the last time.

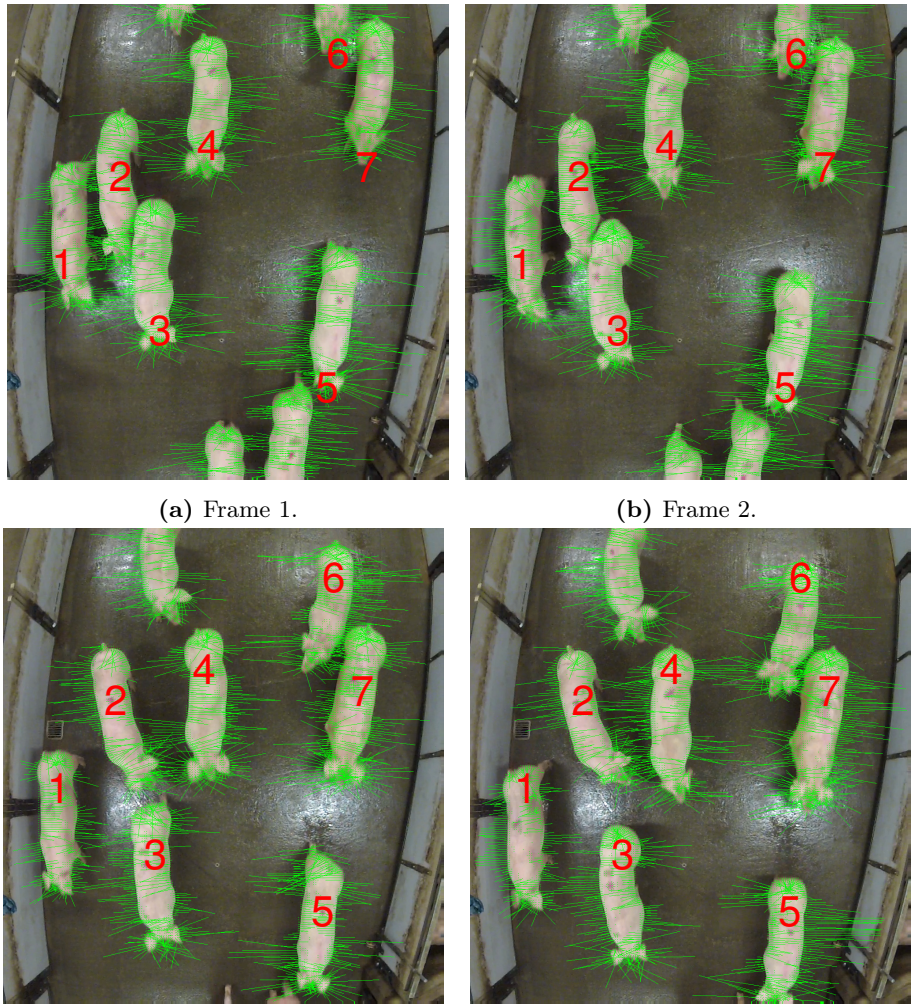
The results of the classification test suggest that lens correction did not affect the correct classification rate of frames without events. However, the method significantly improved the performance classification of frames with events by 23%.

The proposed lens and foreshortening distortion correction method uses measurements of moving objects as references. In the proposed method, the OF vectors are corrected, rather than correcting the image and then estimating the OF. This method could be applied to correct any type of distortion or combination of distortions. The method is also presented and discussed in Paper C.

7.4 Results of stationary pig identification

This section presents the results of the stationary pig identification. A video clip of 150 frames from the *entrance data set* was used for the experiment. At the beginning of the video clip, seven pigs moved from top to bottom in Figures 7.6. However, the second pig stopped. The proposed method (see section 6.6) is used to detect the stationary pig.

For the method training, 382 pigs were individually identified and annotated. 277 pigs were annotated as stationary and 105 as moving. The sample density



(a) Frame 1. **(b)** Frame 2.

(c) Frame 16. **(d)** Frame 26.

Figure 7.6: Indexed pigs in the image are used for further analysis. Green lines represent OF vectors. As can be seen from the figures, pig number 2 did not move forward.

of the testing stage of out-of-control vectors is presented in Figure 7.7. There are some out-of-control vectors over the entire frame, but the highest peak is over the stationary pig.

The proposed method identified the stationary pig in 12 out of 30 frames (see

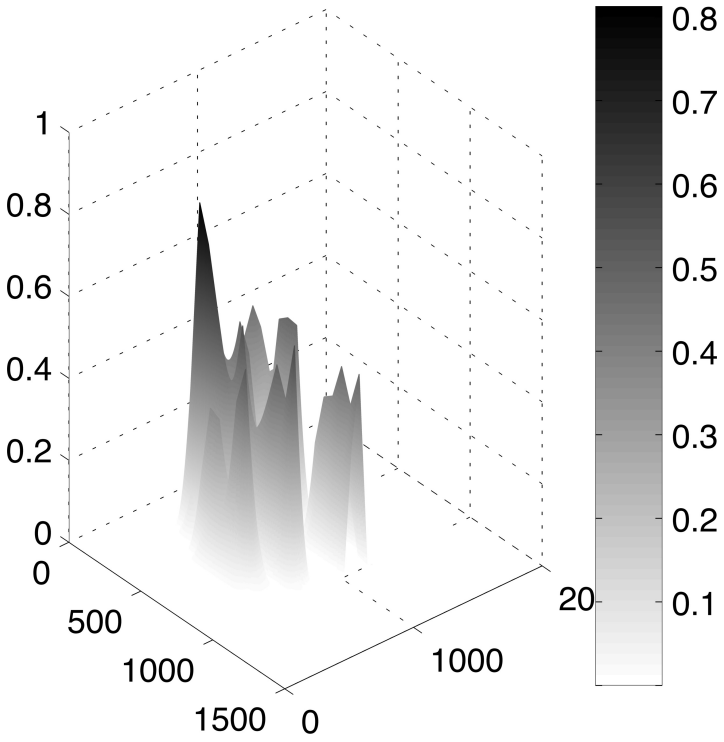


Figure 7.7: Identification of a stationary pig.

Figure 7.8). In 18 frames, the pig did not take any steps, but its body continued moving due to inertia. The same method could be used to identify a pig that is moving too fast, which can indicate stress. The method is also presented in Paper B.

7.5 Motion features of a pig herd

In the video surveillance, motion estimation and object detection are followed by: (1) event or action modeling and (2) classification. In crowd monitoring, features such as density and velocity are often estimated for an event detection. As described in section 6.3, usual measures are not sufficient when surveyed animals are large and have a specific walking pattern, such as a trot. In this section, the efficiency of basic statistics (such as mean, variance, skewness, and kurtosis) of the combined and selected herd OF vectors are analyzed. This follows Dawkins [44, 43], who related these basic statistical summaries to the

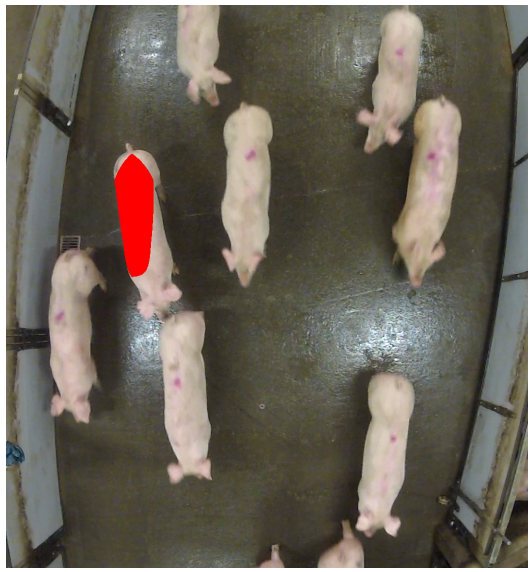


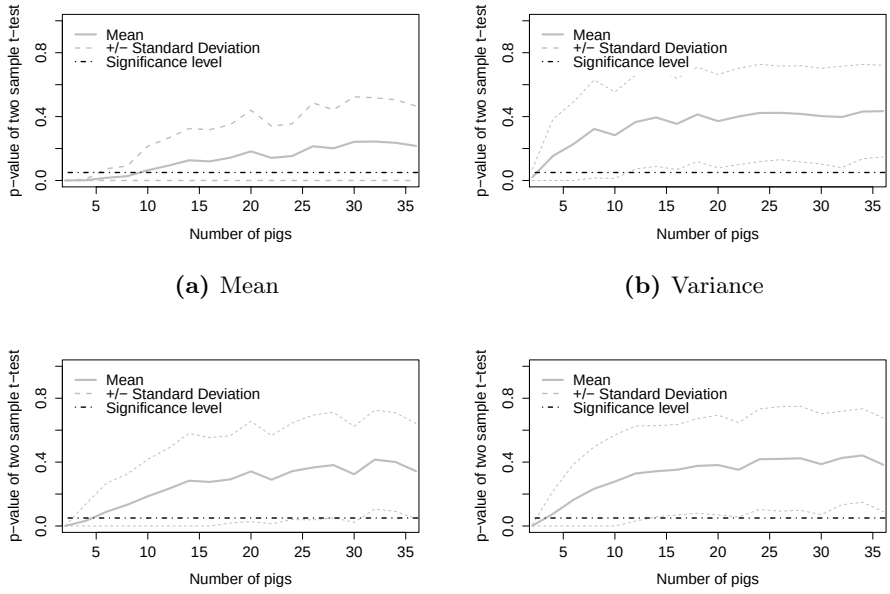
Figure 7.8: Identification of a stationary pig.

welfare score of chicken broilers.

The artificial frames were constructed using sampled pigs from the *entrance data set* in order to evaluate the performance of the low-level statistics proposed by Dawkins. Two groups of frames with different numbers of pigs were constructed. In the first group, all pigs were moving; in the second group, only one pig did not move. The number of pigs in each group ranged from 2 to 36. 30 frames for each group and for each number of pigs were constructed. The low-level statistics were estimated for each frame and a *t-test* was performed to identify if the stationary pig is present in a frame. The test was repeated 500 times. The results in Figure 7.9 show that only mean statistics can be used for motion analysis in a herd of up to 5 pigs.

As suggested in section 6.6, the MAH bins varied in their power to differentiate between moving and stationary pigs. Four vectors, pointing towards the direction of movement, were the most important in the identification of a stationary pig. The test, as described above, was performed to investigate whether the low-level statistics of selected bins could be used to detect stationary pigs. The results in Figure 7.10 show that low-level statistics are not suitable for identifying stationary pigs in a large herd when selected features are used.

The test indicates that the low-level statistics cannot be used to identify abnormal pig behavior. The case study is also presented in Paper B.



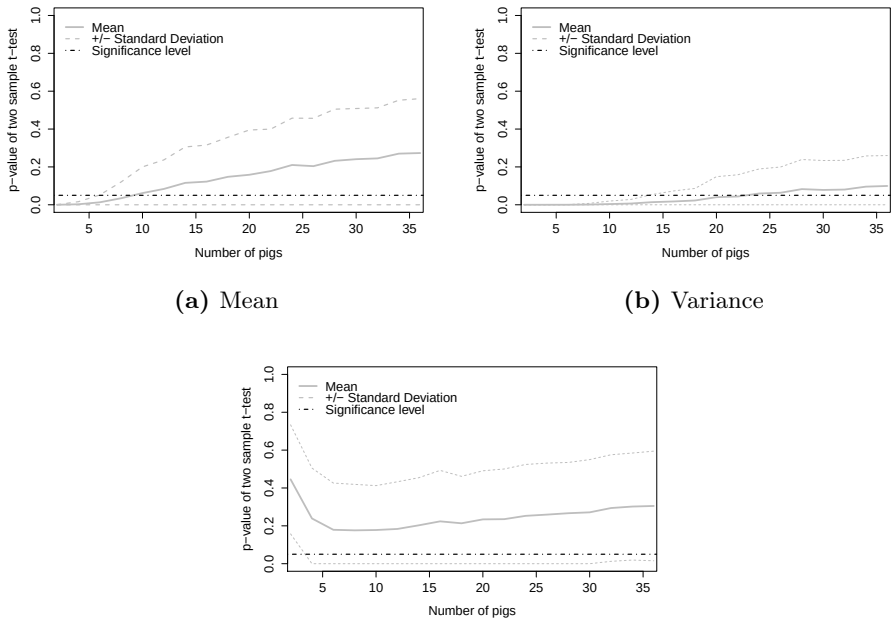
(a) Mean **(b)** Variance

(c) Skewness **(d)** Kurtosis

Figure 7.9: *p-values* of the two-sample *t-tests* for the low-level statistics between the two groups. Solid lines represent mean values, while dashed lines represent ± one standard deviation from the mean.

7.6 Motion classification

The results of the event detection are presented in this section. As described in section 6.1, SVM (see section 3.2) is the most commonly used method for event detection. In this thesis, SVM performs classification based on two sets of features: (1) MAH (see section 6.3) and (2) the relative number of pigs. For illustration purposes, the mean MAHs of four *unloading data set* subgroups are presented in Figure 7.11. The groups were as follows: (1) freely walking pigs; (2) pigs were encouraged, but no event was present; (3) pigs were encouraged, and tripped; and (4) pigs were encouraged and stepped on each other. As can be seen, the subgroups containing events had higher peaks than those without events. The relative pig count is estimated by counting the number of pixels classified as pigs (see section 7.2). The mean density measures for each subgroup are presented in Table 7.1. The subgroups, which contain events have a higher relative count of pigs than those without events.



(a) Mean

(b) Variance

(c) Kurtosis

Figure 7.10: *p-values* of the two-sample *t-tests* for the mean, variance, and kurtosis of an out-of-control area. Solid lines represent mean values, while dashed lines represent ± one standard deviation from the mean.

The variables used for SVM were the average of the two highest bins in the (1) left peak, and (2) right peak, as well as (3) the average of the two smallest bins in the valley, and (4) the number of pixels classified as pigs. The MAH values of the bins were highly correlated. Therefore only the peaks, valley information and relative count of the pigs were used. The data were divided into training sets (75% of the frames) and testing sets (25% of the frames), to identify the optimal SVM scaling parameter σ. In total, 93% of all frames were classified correctly, including 90% of the frames with events, and 94% of the frames without events. A total of 12 separate events were present in the analyzed data set, and the method identified some of the frames from each the events. The results of this analysis are also presented in Paper C.

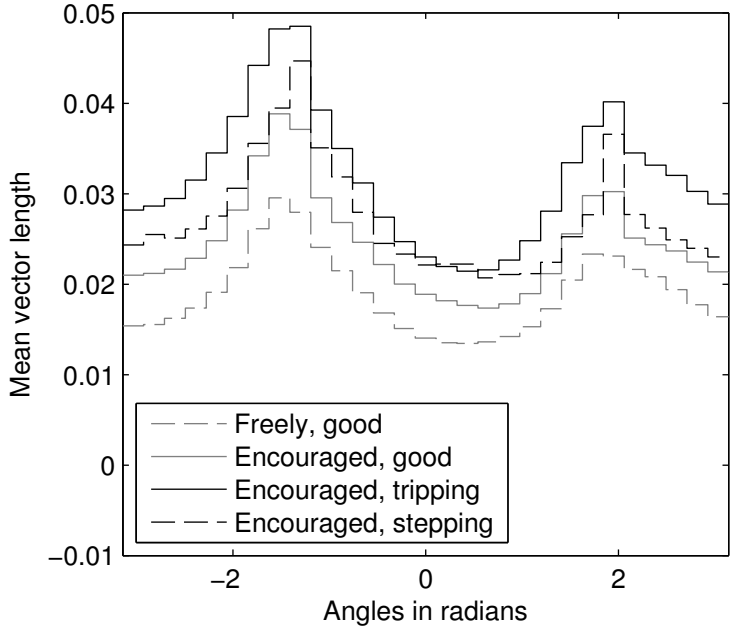


Figure 7.11: Mean MAH of four sub-classes.

7.7 Continuous motion monitoring

The SVM CDV (see in section 6.4) monitored for continuous motion monitoring in this thesis. The CDV were noisy and auto-correlated, and data therefore had to be smoothed. The LOESS was used for smoothing and the CUSUM chart was used for decision making. Three parameters had to be determined in this approach: (1) the smoothing span N; (2) $(M + K)$, and (3) the threshold H. The parameters could be chosen to optimize the (1) total classification rate, (2) early detection rate, or (3) could be based on the SPC approach. The best parameters for the optimization of each criterion are presented in this section. The parameters to be considered were $N = 5, 10, 20, ..., 60$, $(M + K) = 0, 0.05, 0.1, ..., 0.6$, and $H = 0.001, 0.004, ..., 0.01$. For this analysis, 15 clips from *unloading data set* were used. In each clip, one or two sections were annotated. In total, 11 sections were declared as "out-of-control", and 7 as "in-control".

Two sets of parameters were identified to optimize the total classification rate ($N = 45$, $(M + K) = 0.35$, $H = 0.05$ and $N = 50$, $(M + K) = 0.35$, $H = 0.001$) and one set of parameters was identified to optimize the early detection rate ($N = 35$, $(M + K) = 0.6$, $H = 0.001$). The performance of the identified

Movement sub-classes	Number of pixels
moving freely	98403
encouraged but moving ok	135261
encouraged and tripped	150449
encouraged and stepped on each other	208412

Table 7.1: Average pixel counts for different sub-classes.

		Set 1		Set 2		Set 3	
	SVM	CUSUM	p-val	CUSUM	p-val	CUSUM	p-val
Total	95.6 %	99.8 %	0.998	99.8%	0.998	97.7%	0.966
In	99.8 %	100 %	0.977	100%	0.977	96.4 %	0.186
Out	74.3%	99.3%	0.002	99.4%	0.002	99.9%	0.002

Table 7.2: EDCs comparison of three CUSUM parameter sets. For all parameter sets, events are detected earlier compared to the standard approach ($p < 0.05$).

parameters was compared to discrete SVM. A Wilcoxon signed-rank test was used for the comparison. The total, in-control and out-of-control classification rate comparison is presented in table 7.2. The early detection rate comparison is presented in Figure 7.12 and Table 7.3.

The comparison indicates that there was no significant improvement in the total and in-control frame classification rates. SVM correctly classified most in-control frames, meaning that significant improvement was not possible. However, the CUSUM chart significantly improved the out-of-control frame classification with all sets of parameters. The set 3 parameters detected events earlier than set 1 and 2.

Parameters were identified based on the SPC approach as explained later in this

Set	p-value	Number of frames detected earlier than SVM
1	0.006	81.18
2	0.006	81.81
3	0.006	139

Table 7.3: EDCs comparison of three CUSUM parameter sets. For all parameter sets, events are detected earlier compared to the standard approach ($p < 0.05$).

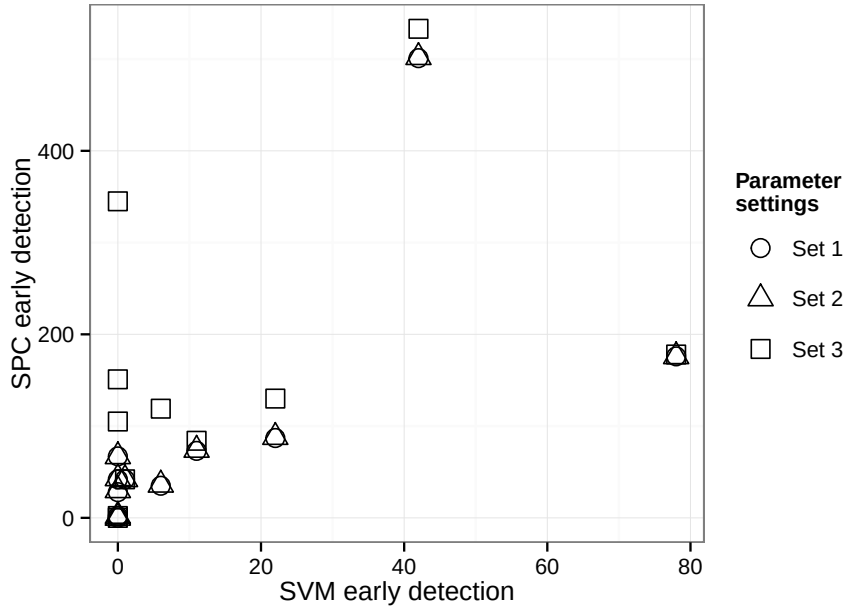
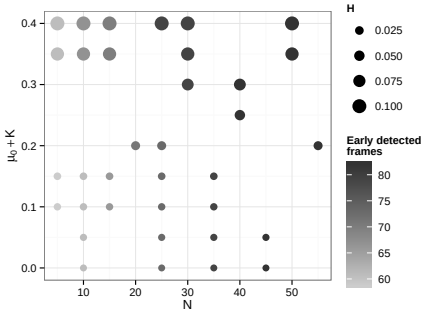


Figure 7.12: EDCs comparison of three parameter sets.

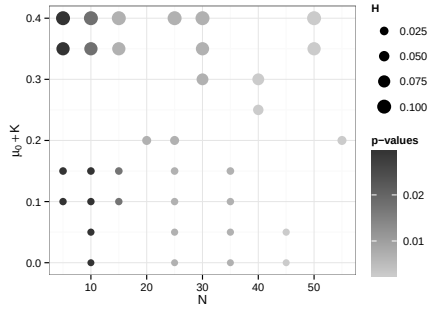
section. The false alarm rate was fixed to the discrete SVM misclassification rate of annotated in-control frames, $\alpha = 0.1705$. For different combinations of N and $(\mu + K)$, a minimum value of H was identified, such that the in-control classification rate of annotated frames using a CUSUM chart was 99.8295%. A total of 35 sets of parameters N, $(\mu + K)$ and H were identified. The early detection rates and the out-of-control classification rates are compared in Figure 7.13.

The CUSUM chart detected events from 59 to 82 frames, or approximately $2 - 2.7$ seconds earlier. The out-of-control classification rates are achieved up to $98\% - 100\%$, which was 25% more than discrete SVM.

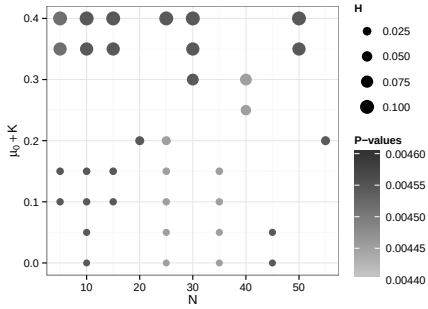
The results presented above indicate that all sets of identified parameters perform better than discrete SVM. The case study is also presented in Paper D.



(a) Mean of early predicted frames.



(b) Wilcoxon test for early predicted frames.



(c) Wilcoxon test for out-of-control frame classification.

Figure 7.13: Comparison of the SVM and CUSUM chart. The p values indicate that a CUSUM chart detects events significantly earlier and correctly classifies significantly more frames with events, at a 0.05 significance level. The CUSUM chart can detect events from 59 to 82 frames earlier than SVM.

7.8 Comparison of two methods for assignable cause identification

Once an undesirable behavior has been identified in a surveyed area, it is also important to find the cause of disturbance. In this section, the two approaches are used to identify assignable causes based on (1) binomial distribution (see section 4.1) and (2) MC (see chapter ch:MarkovChain). A discussion and the results of the binomial distribution approach are also presented in Paper C.

The *unloading data set* was divided into three subsets according to the drivers

that unloaded the trucks, in order to investigate their influence on the behavior of the pigs. Each set was classified using discrete SVM and the number of out-of-control frames was counted. A logistic regression model was built to describe dependent (ratios of out-of-control frames and a total number of frames per driver) and independent variable (drivers). The estimated coefficients are presented in Figure 7.14. A pairwise comparison of the influence of the drivers on the pigs' behavior was performed using ANOVA, and the results are presented in Table 7.4. The p-values indicate that the drivers all had a different influence on behavior at a significance level of 0.05, although Figure 7.14 shows that Drivers 2 and 3 were more similar than Driver 1.

	p-value
Driver 1 vs. 2	0.001
Driver 1 vs. 3	0.001
Driver 2 vs. 3	0.0002352

Table 7.4: P-values for ANOVA test comparing drivers. All the drivers perform differently at significance level $\alpha = 0.05$.

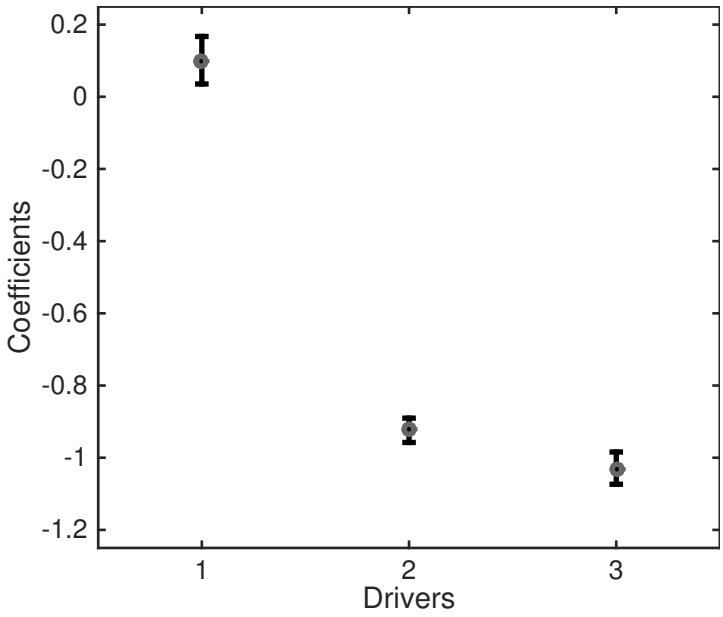
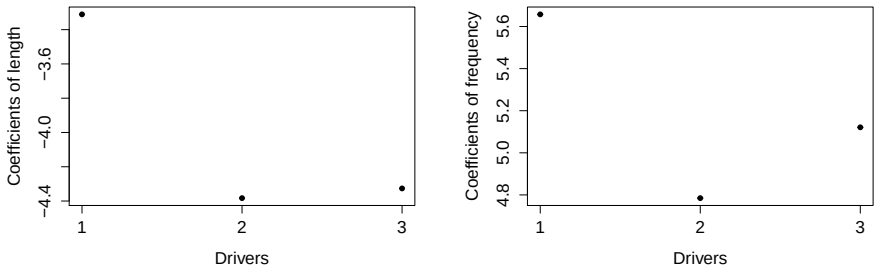


Figure 7.14: Estimated coefficients and confidence intervals for logistic regression model.

In the first step, the transition matrices (see section 6.7) for each driver are

estimated for the performance comparison based on the MC approach. In the second step, the p and $1-q$ probabilities are compared using ANOVA and results indicate that all comparisons are significantly different at significance level 0.05. The coefficients for the logistic regression are presented in Figure 7.15.



(a) Frequency **(b)** Length

Figure 7.15: Estimated coefficients and confidence intervals for logistic regression model.

As with the binomial distribution, the p-values indicated that the drivers all influenced the pigs' behavior to differing degrees, with a significance level of 0.05, although Figure 7.15 shows that Drivers 2 and 3 were more similar than Driver 1. However, using the MC-based approach can give a better understanding of the differences in performance. For example, Driver 3 was 1.06 times more likely to have longer events, and 1.4 times more likely to have more frequent events than Driver 2. Driver 1 was on average 2.84 and 2.06 times more likely to have longer and more frequent events, respectively, than Drivers 2 and 3.

Using both binomial distribution and the MC-based approach in the performance comparison of the three drivers gives a similar overall conclusion. However, a more precise, detailed performance interpretation can be obtained with the MC-based approach.

Chapter 8

Discussion and conclusion

8.1 Discussion

The aim of this thesis is to provide a tool to ensure pig welfare at a slaughterhouse by monitoring the herd behavior using video surveillance. Various methods have been proposed in the literature for tracking pig movement. The methods use markings, features or other techniques for tracking pigs in a constrained environment. However, these methods are not suitable when monitoring a dense pig herd in an unconstrained environment (when pigs can leave and re-enter the surveyed area). Such cases are common in human crowd surveillance, and OF-based approaches, instead of individual tracking, are used. However, these methods cannot be directly applied when monitoring pigs. Pigs have a specific walking pattern (such as a trot), and appear large when the video is recorded from above.

In this thesis, an OF-based approach is proposed to monitor pig herds. There are various methods proposed in the literature to summarize OF, but these approaches are not suitable when monitoring pigs from above. To summarize movement and accommodate a trotting movement, MAH is used. MAH is easy to construct and interpret and there are several advantages of using such an approach. For example, individual pigs do not have to be identified and tracked. Another advantage is that animals can leave and enter the surveyed area without critical information being lost. The third advantage of using MAH together

with SPC is that it can be used to detect undesirable behavior in a herd.

It is common to use radial lens distortion for convenient surveillance. There are several methods proposed in the literature for distortion correction. In this project, additional distortion (foreshortening distortion) was present in the videos. A method for correcting a combination of distortions is proposed. This method corrects OF vectors, rather than a frame. This method is fast and easy to implement, and it can be applied any distortion or combination of distortions can be present in a frame. A disadvantage is that the length and width of multiple objects throughout the frame must be recorded in order to train the method, which is not always possible.

The detection of events or any abnormal behavior in human and animal surveillance is based on discrete classification. The disadvantage of this approach is that it performs in extremes: either the event is present or it is not. Most types of motion are continuous and do not have a clear beginning or end. In this thesis, the principles of SPC were applied for continuous behavior monitoring in order to capture the transitions between events. Most classification methods perform final classification based on continuous decision values such as probabilities, scores, etc. Pig behavior was classified using SVM and continuous SVM decision values instead of discrete SVM. The continuous values were noisy, thus LOESS smoothing was applied, and the CUSUM chart was used for final behavior classification. The suggested approach increases the early detection of events and decreases the false alarm rate. This approach is suitable for monitoring pig behavior as well as human behavior.

It is common to perform a comparison study between different groups to detect assignable causes. Comparison studies often use binomial distribution, which is based on the assumption that events are independent. In the case of pig surveillance, the source of assignable causes can be due to the lack of driver experience, the length of transportation time, or the fattening institutions. The observations from the video are highly auto-correlated, therefore, a new approach must be developed. A method based on MC is proposed in this thesis. This method performs a comparison of the frequency of the event occurrence and the length of the event. The method shows promise, yet more research must be done to validate the approach.

8.2 Conclusion

The increased attention paid to animal welfare and the large numbers of animals being slaughtered every day have created a great interest in an automated pig behavior monitoring system. This thesis aims to provide a tool for monitoring pig behavior using video surveillance to ensure animal welfare. Several data sets

recorded at a slaughterhouse were analyzed. Events such as pigs tripping over and stepping on each other are of interest as they are associated with stress.

The main contribution of this thesis is a proposed framework for monitoring large herds of animals in an unconstrained environment. The work is based on a human surveillance framework. The OF is used to estimate the motion of the pigs. The proposed MAH is used to summarize this motion, and SVM is used to detect abnormal behavior. Continuous SVM decision values (rather than discrete classification) are used in the final decision making. This approach is unique to this thesis and has not previously been proposed. LOESS smoothing is applied to SVM CDV to remove the noise, and the final decision is made based on the CUSUM chart. The use of continuous SVM increases the event detection rate by 25%.

In order to locate pigs moving abnormally within a herd, MAH together with principles of SPC are proposed. This method located a stationary pig in 12 out of 30 frames.

The videos used for the analysis were recorded with radial lens and foreshortening distortion. A method for correcting a combination of distortions is proposed. The method corrects the OF vectors instead of a frame. Three indirect tests indicated that the method performs well.

A new approach is proposed for the analysis of assignable causes of abnormal behavior. The method is based on MC, and future work should focus on validating this approach.

Appendix A

Appendix 1

R. Gronskyte, M. Kulahci, L. K. H.Clemmensen, "Monitoring Motion of Pigs in Thermal Videos", in *Workshop on Farm Animal and Food Quality Imaging*, pp. 31-36, 2013.

Monitoring Motion of Pigs in Thermal Videos

Ruta Gronskyte, Murat Kulahci, and Line Katrine Harder Clemmensen

Technical University of Denmark, DTU Compute,
Richard Petersens Plads 324-220 , 2800 Kgs. Lyngby, Denmark
{rgro,muku,lkhc}@dtu.dk

Abstract. We propose a new approach for monitoring animal movement in thermal videos. The method distinguishes movements as walking in the expected direction from walking in the opposite direction, stopping or lying down. The method utilizes blob detection combined with optical flow to segment the pigs and extract features which characterize a pig's movement (direction and speed). Subsequently a multiway principal component analysis is used to analyze the movement features and monitor their development over time. Results are presented in the form of quality control charts of the principal components. The method works on-line with pre-training.

Keywords: Optical flow, blob detection, multiway principle components, quality control.

1 Introduction

Animal well-being has become a concern for consumers and [1] suggests that the stress level of pigs before slaughter influences meat quality. To ensure animal well-being the pigs should be constantly monitored and in case of a stressful situation actions should be taken. However it is difficult to keep track of many animals and therefore some automated behavior analysis methods should be implemented. For this paper, pigs were filmed in a constrained area walking from left to right. However, some pigs can change direction or stop walking. Such events can block the movement of other pigs. There can be different reasons for the change in movements such as not feeling good or an obstacle appeared in the path. The classification is challenging, because it is quite normal for pigs to slow down or even stop to sniff for no reason but out of curiosity.
The automated video analysis will allow the slaughter house to make sure all animals are walking in order and intervene when necessary. It is important, that the analysis provides a fast overview of the area with easily interpretable results. No animal crowd monitoring and analysis methods have been suggested in the literature. Previous research has mainly focused on analyzing human crowd behavior in surveillance videos. A good overview of the methods can be found in [2]. The choice of method greatly depends on the video type and what we are looking for in the videos. There are methods available for tracking individual objects, usually used for pattern search in movements. However, in our thermal

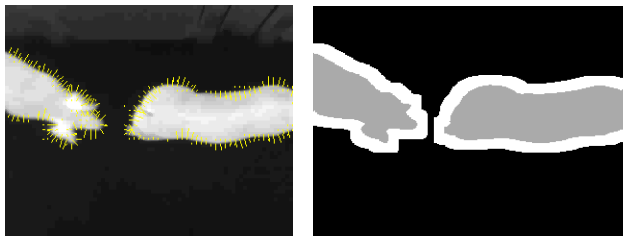
videos it is very complicated to identify the individual pigs because of physical similarities and the fact that each pig does not necessarily appear in many frames. Therefore we instead propose to use optical flow which often is used for object tracking and action recognition. This method gives a great overview of the surveillance area.

2 Methodology

In this section the methodology is presented in details. It takes two distinct steps to perform the analysis. In the first step, the visual analysis is performed using optical flow, blob detection and optical flow quantification. The second step is the behavioral analysis based on quality control charts. Here multiway PCA is performed and quality control chats are built for the principal components.
We used different sections from 5 thermal videos. In total 2460 frames were available for training. For testing representative sections from 2 thermal videos were extracted with a total of 2284 frames. To validate the test results the 2284 frames were manually annotated and classified.

2.1 Visual Analysis

As mentioned above we are not just interested in detecting moving pigs but also the stationary ones. To do so we merged two methods: optical flow and blob detection. First optical flow is applied and then filtered by a simple threshold to remove the noise. The threshold is half of the overall average length of the vectors from optical flow. The results of this step for one frame are shown in Figure 1.



(a) Optical flow. (b) Blob detection.

Fig. 1: Visual analysis step. First we calculate optical flow and then use blob detection. In (b) grey represents the actual blobs and white represents blobs extended by 5 pixels.

To separate those optical flow vectors representing pigs from the background we created a binary mask using morphological erosion and opening. These are

particularly convenient as both are obtained as by-products of optical flow. Alternatively a simple threshold could be used. All blobs were extended by 5 pixels to include the vectors along the edges in the further analysis.
For each frame two histograms were used to quantify optical flow. The first represents the lengths of the optical flow vectors and the second the angles. The number of bins were selected by

2.2 Quality Control

Multiway PCA is used in batch monitoring in statistical process control[3] . Investigating the quality of a current batch requires historical data of good batches. Data consist of repeated measurements monitored throughout the process. A collection of batches can be presented in 3D matrix and a special unfolding technique to a 2D matrix will allow to apply ordinary PCA. By monitoring the score plots of principal components it is possible to track changes in the process. For multiway PCA application on thermal videos we need to define what we mean with "the batch". We use the concept of a scene: a constant number of consecutive frames in a video is a scene. The number of frames per scene was found by minimizing the prediction sum of squared residuals (SSE) on a training set including all PC.

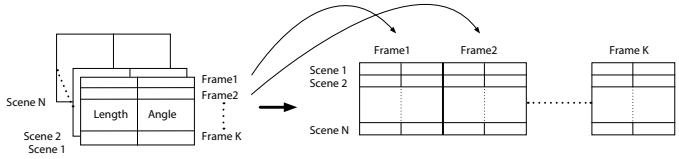


Fig. 2: Unfolding the data matrix.

As it was mentioned above a special unfolding technique has to be performed such that the ordinary PCA can be applied. Let$N$ be the number of scenes and K the number of frames in each scene. Each frame is represented by the counts from the two histograms which are stacked next to each other. The unfolding is done by reshaping the scene to a row vector, i.e. the K frames of a scene are stacked after each other as shown in Figure 2. All the unfolded scene vectors are stacked on top of each other forming the final matrix. Let J be the total number of bins per frame, then the unfolded matrix has the dimension $N \times JK$. This unfolding technique allows for comparison among scenes.
A score matrix t, loading matrix p and residual matrix E were obtained after performing PCA on the unfolded matrix. R is the number of principal components. Let X be unfolded matrix then it can be presented as:

$$X = \sum_{r=1}^{R} t_r \otimes p_r + E \tag{1}$$

In statistical quality control a training state is usually called phase I. In this phase we collect good scenes, build a quality control chart and check if all our scenes are statistically in control. The control limits used in this phase are different from the limits used in the second phase. In [4] they suggest three methods for checking good batches. First Hotelling's T2 statistics:

$$D_s = t'_R S^{-1} t_R \frac{I}{(I-1)^2} \sim B_{\frac{R}{2},\frac{I-R-1}{2},\alpha} \tag{2}$$

where $S \in \mathbb{R}^{R \times R}$ is an estimated covariance matrix and B is a beta distributed random variable. The second test is a sum of square of residuals of individual batches:

$$Q_i = \sum_{k=1}^{K} \sum_{j=1}^{J} E(i, kj)^2 \tag{3}$$

For the third test the PCA scores are used. Score plot of the first two principal components and confidence internals are used to identify outliers. The confidence intervals are ellipsoids with center at $\mathbf{0}$ and axis length:

$$\pm S(r,r) B_{1,\frac{I-2-1}{2},\alpha} \sqrt{\frac{(I-1)^2}{I}} \tag{4}$$

In phase II we perform on-line monitoring. For the on-line monitoring new confidence intervals for the score plot must be calculated:

$$\pm S(r,r) F_{2,I-2,\alpha} 2 \sqrt{\frac{I^2-1}{I(I-2)}} \tag{5}$$

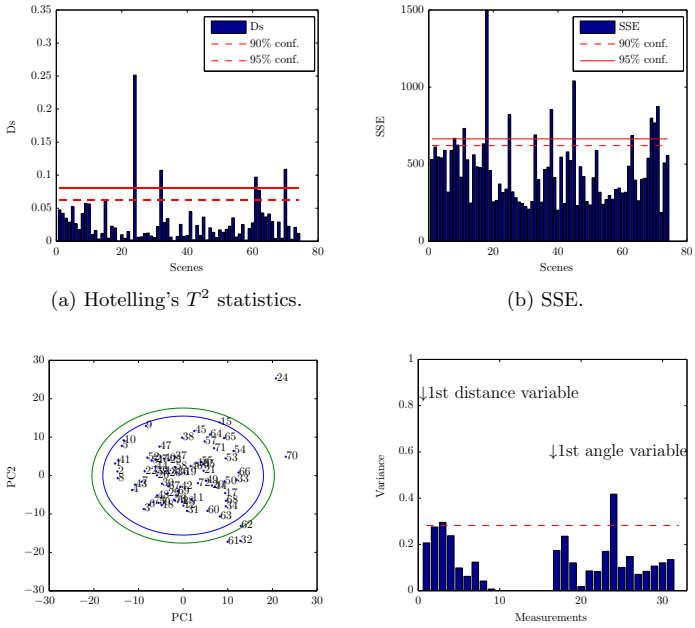
A visual analysis was done for every frame when on-line monitoring had started. Every set of 25 frames form a scene which is transformed into a score through the multiway PCA. The score is added to the quality control chart. [3] suggests not waiting for all measurements from a batch but to estimate the remaining batch measurements. However, there is no reason to do so here since a scene only requires 25 frames, thus control chart is updated every few seconds.

3 Results

As mentioned above, two phases are required to perform the analysis of thermal videos. In this section results of each phase will be discussed.

3.1 Phase I

Figure 3 shows Hotelling's T^2 statistics (a) and SSE (b) for every scene, and the scores of the two first principle components (c). The first two principal components were chosen naively as Hotelling's T^2 statistics combines the PCs equally weighted causing increased misclassification when including additional



(a) Hotelling's T^2 statistics.

(b) SSE.

(c) Score plot.

(d) Explained variance by each variable.

Fig. 3: Training data.

components. Analyzing many plot is not an option as well, because the aim is to give an easy to interpret overview of the video. These three plots all have points exceeding the confidence interval thus indicating that there might be some outliers. However, after inspecting each scene no unusual behavior was noticed. Figure 3(d) shows the explained variance by each of the 32 variables. The most important variable is the 8th variable from the angle histogram. This bin represents vectors with the smallest angles. A small angle is when pig is walking straight. The second most important variable is the 3rd bin of speed. The faster the pigs are going the heavier the tail of the speed histogram will be.

3.2 Phase II

Each of the 2284 frames were manually annotated as not moving if at least one pig was not moving. A scene was declared as not moving if more than half of the frames were annotated as not moving. Table 1 shows that 66% of all scenes were classified correctly and at the individual frame level 78% of all frames were

classified correctly. As it can be seen in Figure 1 most of the errors appeared very close to the limits. It is important to remember, that it is very difficult to annotate movements just by looking at a single frame or even a sequence of frames. Some errors could appear due to annotation.

Annotated \ Classified	Moving	Not moving
Moving	17	8
Not moving	21	36

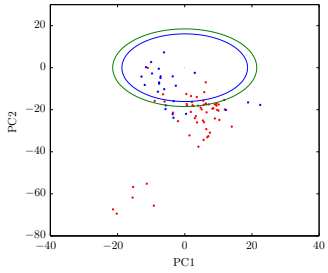


Table 1: Results of phase II.

4 Conclusion

Our suggested method can classify 66% of scenes and 78% of the frames correctly. It is difficult to get higher results due to the complexity of annotation. Also some pigs may slow down to sniff around but this situation should not be considered as not moving. However, these situations will create additional variance.
Future improvements could be to analyze clusters or individual pigs and new methods for vector quantification. In scenes with many pigs and lots of action some details can get lost in the histograms.
With better quantification of the optical flow vectors it would be possible to determine some patterns of behavior or actions through classification based on score plots.

References

1. P.D. Warriss, S.N. Brown, S.J.M. Adams, and I.K. Corlett, *Relationships between subjective and objective assessments of stress at slaughter and meat quality in pigs*, Meat Science **38** (1994), no. 2, 329–340.
2. Weiming Hu, Tieniu Tan, Liang Wang, and S. Maybank. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 34(3):334–352, 2004.
3. Paul Nomikos and John F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):13611375, 1994.
4. Paul Nomikos and John F. MacGregor. Multivariate SPC charts for monitoring batch processes. *Technometrics*, 37(1):41, February 1995.

Appendix B

Appendix 2

R. Gronskyte, L. H. Clemmensen, M. S. Hviid, M. Kulahci, "Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms", *Biosystems Engineering*, vol. 141, pp. 19-30, 2016.

BIOSYSTEMS ENGINEERING 141 (2016) 19–30

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms

Ruta Gronskyte [a,*], Line Harder Clemmensen [a], Marchen Sonja Hviid [b], Murat Kulahci [a,c]

[a] DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark
[b] Danish Meat Research Institute, Taastrup, Denmark
[c] Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form
16 August 2015
Accepted 5 October 2015
Published online xxx

Keywords:
Optical flow
Support Vector Machines
Modified angular histograms
Abnormal movement detection

We analyse the movement of pig herds through video recordings at a slaughterhouse by using statistical analysis of optical flow (OF) patterns. Unlike the previous attempts to analyse pig movement, no markers, trackers nor identification of individual pigs are needed. Our method handles the analysis of unconstrained areas where pigs are constantly entering and leaving. The goal is to improve animal welfare by real-time prediction of abnormal behaviour through proper interventions. The aim of this study is to identify any stationary pig, which can be an indicator of an injury or an obstacle. In this study, we use the OF vectors to describe points of movement on all pigs and thereby analyse the herd movement. Subsequently, the OF vectors are used to identify abnormal movements of individual pigs. The OF vectors, obtained from the pigs, point in multiple directions rather than in one movement direction. To accommodate the multiple directions of the OF vectors, we propose to quantify OF using a summation of the vectors into bins according to their angles, which we call modified angular histograms. Sequential feature selection is used to select angle ranges, which identify pigs that are moving abnormally in the herd. The vector lengths from the selected angle ranges are compared to the corresponding median, 25th and 75th percentiles from a training set, which contains only normally moving pigs. We show that the method is capable of locating stationary pigs in the recordings regardless of the number of pigs in the frame.

1. Introduction

Today many consumers are increasingly interested in the welfare of the animals used for commercial meat production. Animal behaviour during the production process can be used to evaluate animal welfare (Brandt, Rousing, Herskin, & Aaslyng, 2013). However, constant monitoring of animals by humans in industrial farms is nearly impossible. Such constraint has created great interest in automated livestock

* *Corresponding author*. Tel.: +45 45 25 53 51.
E-mail address: rgro@dtu.dk (R. Gronskyte).
http://dx.doi.org/10.1016/j.biosystemseng.2015.10.002
1537-5110/

monitoring. There have been several attempts to identify (Tu, Karstoft, Pedersen, & Jørgensen, 2013; Guo, Zhu, Jiao, Ma, & Yang, 2015) and track pigs (Oczaka et al., 2014; McFarlane & Schofield, 1995; Kashiha et al., 2013a; Lind, Vinther, Hemmingsen, & Hansen, 2005; Ahrendt, Gregersen, & Karstoft, 2011) as well as chickens (Dawkins, Cain, & Roberts, 2012; Kashiha, Pluk, Bahr, Vranken, & Berckmans, 2013b; Nakarmi, Tang, & Xin, 2014), and cows (Huhtala, Suhonen, Mäkelä, Hakojärvi, & Ahokas, 2007; Porto, Arcidiacono, Anguzza, & Cascone, 2015) in farms. Different movement and density measures have proved useful in ensuring animal welfare (Rushen, Chapinal, & Passill, 2012; Dawkins et al., 2012; Kashiha et al., 2013b; Youssef, Exadaktylos, & Berckmans, 2015; Nakarmi et al., 2014; Porto et al., 2015). The animals can be tracked either by using radio tags attached for example to the animal's ear (Ng, Leong, Hall, & Cole, 2005; Tøgersen, Skjøth, Munksgaard, & Højsgaard, 2010; Ruiz-Garcia & Lunadei, 2011; Porto, Arcidiacono, Giummarra, Anguzza, & Cascone, 2014), passive transponders injected into the animal's body (Prola, Perona, Tursi, & Mussa, 2010; Caja et al., 2005) or by video surveillance. A computer vision approach is non-intrusive and can be adapted to different animals. The cases mentioned above, which all employ computer vision, monitor animals in farms where all areas of interest can be covered with one or a few cameras. When the videos are recorded in constrained areas, which the animals cannot leave, additional markers (Kashiha et al., 2013a) or features (Ahrendt et al., 2011) can be used to track the animals' movement. Ahrendt et al. (2011) propose a method to track individual pigs and can follow three pigs in a constrained area over an 8-min period without losing track. However, monitoring pigs in slaughterhouses creates additional challenges: (1) large numbers of animals, which are very similar to each other, are often present in small areas, and (2) even multiple cameras cannot cover all of the process line, thus animals can leave and come back into a monitored area. The disadvantage of using the marking technique proposed by Kashiha et al. (2013a), which consists of stamping a specific pattern on the back of the pigs, is that additional people would be needed as well as it being time-consuming to uniquely mark 12,400 (Danish Crown, 2014) pigs per day. Tracking individual features as proposed by Ahrendt et al. (2011) or using tags as proposed by Ng et al. (2005), Prola et al. (2010) and Caja et al. (2005) would be too computationally expensive for a real-time monitoring system. These disadvantages can be overcome by using our proposed technique and by analysing all pigs as a herd with the caveat that the ability to track an individual pig is lost.

In dense crowd surveillance of humans, it is rarely attempted to identify the individual subjects. In most of these cases, optical flow (OF) is used (Helbing, Johansson, & Al-Abideen, 2007; Andrade, Blunsden, & Fisher, 2006; Ali & Shah, 2008). OF is a pattern that represents relative motion between two consecutive frames in a video and is presented as a vector field of motion over the entire frame (Wedel & Cremers, 2011). Each OF vector indicates the direction and the distance of movement of a single pixel. OF estimation is a non-intrusive and computationally cheap method. A standard consumer camera is sufficient to record the movement and can be mounted such that it does not interfere with the usual routine of the subjects. This paper describes a case study for monitoring the welfare of pigs while unloading from trucks at a slaughterhouse that can handle up to 62,000 (Danish Crown, 2014) pigs per week. The majority of pigs in our recordings are crossbreeds between Duroc, Danish Landrace and Yorkshire (Dx(LxY)). Pure Duroc is typically red in colour. Thus, some crossbred pigs are coloured. Our recordings mainly contain pigs that are white with only a few coloured pigs. The colour similarities make the separation of the nearby pigs computationally expensive. The similarities in the body size also make it challenging to track individual pigs. Pigs weighed 100–110 kg and were approximately 6 months of age.

To the best of our knowledge, only a few attempts have been made in the literature to monitor livestock using OF (Dawkins et al., 2012; Dawkins, Lee, Waitt, & Roberts, 2009). To quantify the behaviour of an individual animal, Dawkins et al. (2012) used low-level statistics such as the mean, variance, skewness and kurtosis of the OF vectors of chickens and found a strong correlation between statistics and the Bristol Gait Score, which assesses leg weakness. In this paper, we show that the low-level statistics used by Dawkins et al. (2012) are not optimal when the animals are recorded in proximity and have a specific walking pattern such as trot. In addition, low-level statistics are not suitable for monitoring large numbers of pigs as they average out the few pigs that are moving abnormally. We refer to a specific walking pattern (trot) as "local movement". We define "global movement" of an animal as the overall speed and direction of the animal. OF vectors of monitored pigs show not only the global movement but also the local movement of the animals. Low-level statistics of the OF vectors do not represent the global movement of the animals when a strong local movement, i.e. trot, is present. In this paper, we propose to use modified angular histograms (MAH), which summarise the OF vector lengths within the corresponding angle range. By using MAH, we can filter the local movement out of the OF vectors, leaving only the global movement for further analysis.

Two abnormalities are of interest: (1) pigs moving too quickly and (2) pigs moving too slowly or being stationary. The former can indicate stress as it may happen when animals are encouraged to move too quickly or are agitated by external disturbances. The latter may indicate that an animal is injured or sick, or that there is an obstacle in its path. Identifying abnormally moving pigs will allow for possible interventions to ensure the welfare of the animals. The suggested method can be used for real-time monitoring although additional research is needed to account for the correlation between MAH from consecutive frames.

2. Methods and materials

2.1. Data

In this study, pigs were video-recorded in a slaughterhouse just after unloading from trucks. The daily process was recorded to capture the normal pig movement. A GoPro HERO2

(©2013 Woodman Labs, Inc, California, USA) camera was used for recording and placed so that the optical axis of camera was perpendicular to the pigs' moving path. The frame rate is approximately 30 frames per second, and the length of the videos is 17 min 35 s. The frame size is 1920 pixels in height and 1080 pixels in width. There is lens distortion present in the recording.

The distortion is corrected by adjusting for the size of the pigs, which is observed to be approximately constant (Gronskyte, Clemmensen, Hviid, & Kulahci, 2015). A 3rd order polynomial is used to model the length and width of the pigs and correct for distortion:

$$Z(x,y)=\sum_{j=1}^{4}\sum_{i=1}^{4}p_{ij}x^{4-j}y^{4-i} \quad (1)$$

where Z is the length or width of the pig, x and y are the vector positions in the frame and the p_{ij} is the estimated polynomial coefficient. The effect of the lens correction is presented in Fig. 1. The strongest correction is around the lens focal point.

Five sets of images are used for the analysis: (*set A*) six frames sequence of a single pig moving, (*set B*) six frames sequence with six pigs moving and one stationary pig, (*set C*) 190 frames with multiple pigs where some of them are stationary, and (*set D*) eight frames sequence with 29 pigs with two of them stationary and the rest moving. The experimental set (*set E*) consists of 150 frames sequence, of which 30 frames show one stationary pig. The *sets A, B, D* were used for illustration purposes, while *set C* was used for the final method training and *set E* for the method testing.

2.2. Overview of the method

There are four steps in the training method (1) OF estimation and filtering for each frame in the video, (2) MAH estimation based on the OF for fully visible pigs in each frame, (3) selection of the relevant angles in the MAH, and (4) estimation of the median, 25th, and 75th percentiles of these for normally behaving pigs. The testing part, which can be developed to be real-time, includes five steps: (1) OF estimation and filtering, (2) MAH estimation, (3) comparison of the selected bins to the estimated percentiles, (4) interpolation and (5) thresholding. The summaries of the training and testing procedures are presented in Figs. 2 and 3.

2.3. Optical flow

Tracking individual pigs is a complicated task due to similarities of their body shape, therefore OF is used instead. The OF of the pigs is determined using the Vision Toolbox in MATLAB® (Matlab, 2012a), which uses the Horn-Schunck (Horn & Schunck, 1981) method with a five frame delay. The OF vectors span over an entire frame and subsequently the relevant vectors are filtered to identify only those that represent the pigs. Currently, the OF estimation is off-line and performed using implemented functions in MATLAB®. Thus more development is needed to make the method feasible for on-line applications. For our off-line version, the OF for a single frame requires an average of 0.2669 ± 0.0082 s CPU to compute.

Filtering the OF vectors is performed in two steps. First, a colour decorrelation stretch (Matlab, 2012b) is used to enhance the red colour of the pigs. The decorrelation stretch removes the correlation between the RGB colour channels, thus enhancing colours. Secondly, blob detection, which is a by-product of the OF estimation, (Matlab, 2012a) is used to separate the pigs from the background. Blob detection is a computer vision method that detects regions that differ in properties. Although Tu et al. (2013) and Guo, Zhu, Jiao, and Chen (2014) present some elaborate methods for the identification of pigs which can handle illumination issues, our simple pig identification approach works well in our case as no significant illumination or area changes are present in the analysed recordings.

2.4. Dawkin's low-level statistics

Dawkins et al. (2012) analysed the mean, variance, skewness and kurtosis of the OF vector lengths and showed high correlations to welfare measures. Mean (μ) and variance (σ^2) are the classic probability measures for the expected value (E[·]) and the spread of the data around the expected value of the random variable X. Skewness is a measure of asymmetry of a probability distribution and is defined as:

$$\gamma_1=E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \quad (2)$$

Kurtosis measures the peakedness of a probability distribution and is defined as:

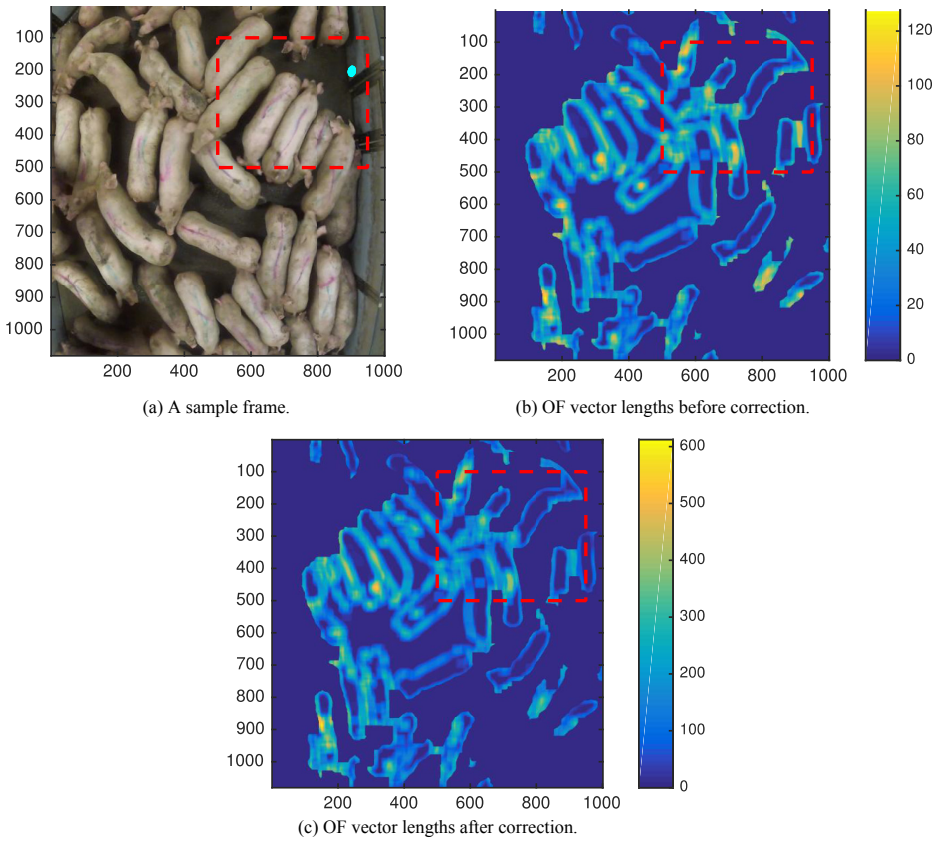
$$\beta_2=\frac{E\left[(X-\mu)^4\right]}{\left(E\left[(X-\mu)^2\right]\right)^2} \quad (3)$$

2.5. Modified angular histogram

To quantify filtered OF vectors we span the vector angle range $[-\pi;\pi]$ and divide it into a constant number of bins. A median length of vectors is estimated within each bin of a corresponding vector angle range. We call this a modified angular histogram. A detailed motivation and description of the MAH are given below.

2.5.1. Motivation for the modified angular histogram

In this subsection, the movement of a single pig from data *set A* (six frames of a single pig moving) is analysed. The elongated body shape and asymmetric front and back leg steps make a pig's movement resemble the letter "S" when observed from above, thus resulting in the OF vectors pointing in multiple directions. Figure 4a shows the OF vectors of a single pig moving from the top of the figure to the bottom. The forward movement corresponds to the angle range between $\left[\frac{\pi}{4};\frac{3\pi}{4}\right]$. The angles of the OF vectors presented in Fig. 4a are symmetrically distributed around a rotational symmetry axis in Fig. 4b. The symmetry axis is diagonal to the pig's body. Patterns of the local movements are detected by OF, and thus a wide range of angles is registered. As a result, not all OF vectors represent the global direction and speed of the pig.



(a) A sample frame.

(b) OF vector lengths before correction.

(c) OF vector lengths after correction.

Fig. 1 – OF vector lengths before (b) and after (c) correction. The dot in (a) is the lens focal point where the correction is strongest. The area marked by the red square is most affected by the correction.

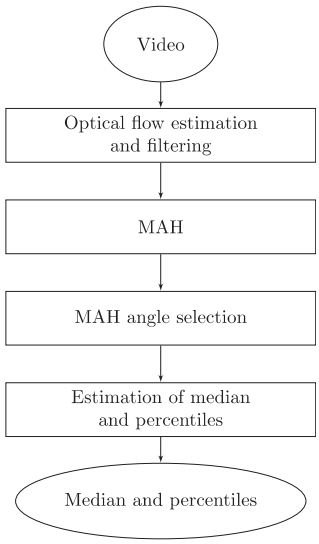


Fig. 2 – Illustration of the training procedure.

2.5.2. *Estimation of modified angular histogram*

We propose to use MAH to quantify and summarise the optical flow vectors of each frame. The idea is to summarise OF vectors with similar angles, and thereby obtain a median direction for each part of the pig. To do this, the angle range $[-\pi; \pi]$ needs to be divided into a set of bins, as in a histogram. We use Doane's formula (Doane, 1976) to obtain the number of bins:

$$k = 1 + \log_2(n) + \log_2\left(\frac{|g_1|}{\sigma_{g_1}}\right) \tag{4}$$

$$\sigma_{g_1} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} \tag{5}$$

where g_1 is the estimated skewness, n is the number of observations, k is the number of bins. For our data, we obtain 31 bins. The median lengths of all OF vectors that fall into the given angle range are calculated for each bin. The median is chosen instead of the mean, due to the mean's sensitivity to outliers.

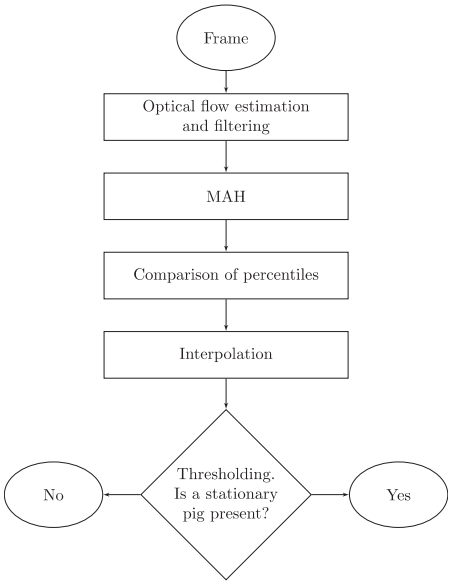


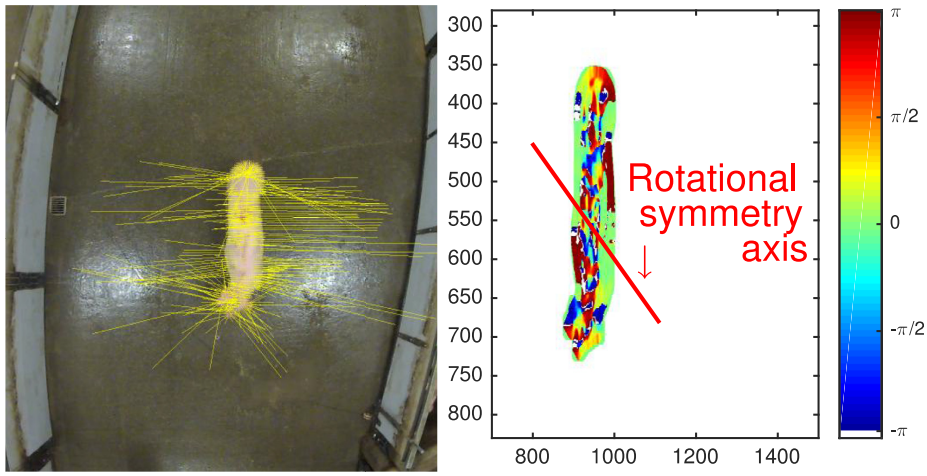
Fig. 3 – Illustration of the testing procedure.

2.5.3. *Feature selection*

Due to a high correlation between the MAH bins, not all bins have to be monitored to identify moving and stationary pigs. A simple feature filtering approach, together with sequential feature selection using the Support Vector Machines (SVM) algorithm is used to identify the most important bins. SVM is a non-parametric supervised classification method that constructs a hyperplane in higher dimensions for classification. In our case, we use SVM to separate moving pigs from stationary pigs. A detailed description of the method can be found in Hastie et al. (2009). Sequential feature selection is widely used to select the most important features within a large set of variables that are highly correlated. The method selects a subset of features by sequentially adding more features until a minimal cross-validation misclassification error is reached. A simple feature filtering is used to help sequential filtering to converge. It is done by performing a *t-test* of the difference in means between moving and stationary pigs on each feature and then comparing the *p*-values.

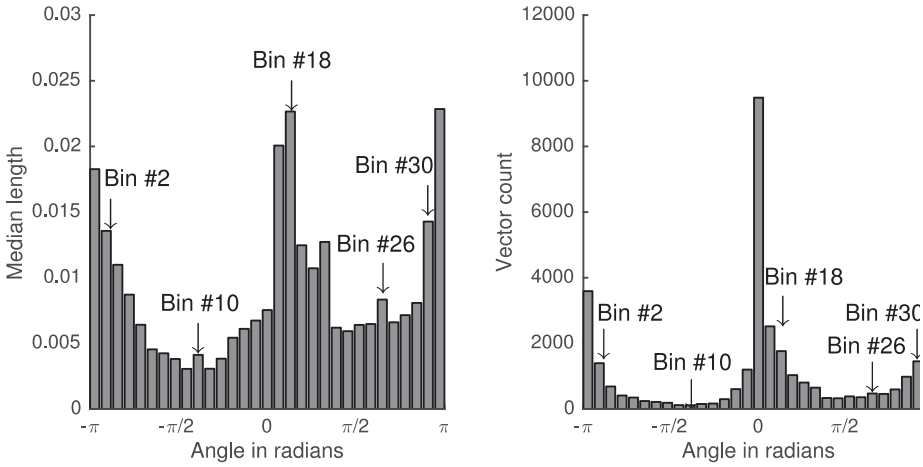
2.6. *Statistical process control*

Statistical process control (SPC) is an approach to monitor and control the process and to ensure the process output is conforming to the specification. The monitoring is performed using control charts that were originally introduced by Shewhart (Montgomery, 2007, p.207–265). For variables of interests that are expected to vary around a constant level, the chart consists of a central line, which is for example the mean or target value of the variable for a good process, as well as upper and lower control limits (UCL, LCL). The process is considered "in control" when the monitored variable is within the pre-defined control limits. Otherwise, it is deemed "out-of-control". Traditionally the main focus has been the central tendency (mean) of the process. Three standard deviations around the mean are used as control limits. The mean and standard deviation are estimated using historical, in-control process data. As mentioned in Section 2.5.2, the median, 25th, and 75th percentiles are employed instead of the mean and standard deviation in the analysed case to decrease the influence of errors accrued in the estimation of the OF vectors.



(a) All identified OF vectors (b) Angles of identified OF vectors.

Fig. 4 – Identified OF vectors for the moving pig and rotational symmetry axis for OF vector angles in radians. The angle range $\left[-\frac{\pi}{4};\frac{\pi}{4}\right]$ represents vectors pointing to the right, $\left[-\pi;-\frac{3\pi}{4}\right]$ and $\left[\frac{3\pi}{4};\pi\right]$ vectors pointing to the left, $\left[-\frac{3\pi}{4};-\frac{\pi}{4}\right]$ vectors pointing backwards, $\left[\frac{\pi}{4};\frac{3\pi}{4}\right]$ vectors pointing forward.



(a) The MAH of vector median lengths.

(b) The MAH of vector counts.

Fig. 5 – The MAHs of the pig presented in Fig. 4.

3. Results

3.1. *Interpretation of MAH*

The MAH of the pig in Fig. 4 is presented in Fig. 5a. Two peaks and two valleys are present in the MAH. The same pattern is present in the vector count histogram in Fig. 5b. The middle peak in the histograms represents a movement to the left, and the two peaks on the sides represent a movement to the right. In a circular plot, the two peaks on the sides would have merged to one peak. The rightmost valley represents the vectors in the global movement direction, and the leftmost valley represents vectors in the opposite direction of the global movement. An overview of the vector lengths and angles is presented in Fig. 6.

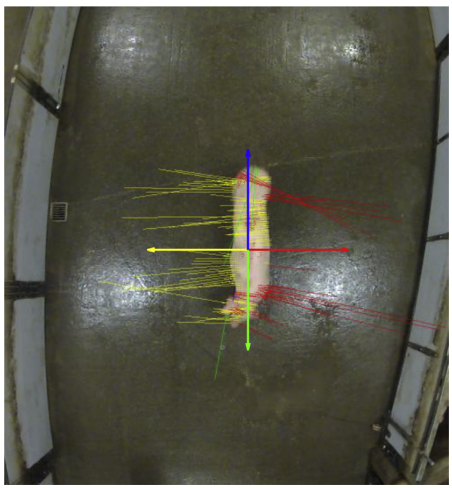


Fig. 6 – The lines represent the 150 times enlarged OF vectors of a single pig, and the arrows represent the directions. Red colour represents vectors that fall in the middle peak of the MAH, yellow vectors fall in the side peaks, blue vectors fall in the left valley and green vectors fall in the right valley.

The MAHs verify that the OF vectors capture not only the forward motion of the pig but also the local movements of its body. The vectors in the global movement direction are fewer and smaller in length than vectors representing local movement; indicating that the body of the pig wiggles rapidly from side to side while it slowly moves forward.

3.2. *Performance of Dawkins low-level statistics on large animals*

From data *set D* (eight frames with 29 moving and two stationary pigs), OF vectors of all pigs are extracted, and all pigs are labelled on the basis of manual assessment of the video sequence as moving or stationary. By sampling from these pigs at random with replacement, two sets of artificial frames are made consisting of the corresponding OF vectors. An artificial frame is a collection of one or several pigs' optical flow vectors, disregarding the location in time and space of the pigs. In the first set, only moving pigs are sampled, and in the second set one stationary pig and otherwise moving pigs are sampled. 30 artificial frames are sampled per set with the same total number of pigs in both sets. We varied the number of pigs from 2 to 36 with an increment of 2 and for each number of pigs we repeat the experiment 500 times. For each frame, the mean, variance, skewness and kurtosis of the length of all OF vectors are estimated. Two-sample *t-tests* for

the difference in means between the two groups are performed for each of the low-level statistics with an increasing number of pigs per frame. A summary of the *p*-values of the two-sample *t-tests* is presented in Fig. 7.

When the number of pigs is large, there is no difference between frames with and without stationary pigs for the low-level statistics. For variance (see Fig. 7b) and kurtosis (see Fig. 7d), there are no differences between the two groups, whereas for the mean (see Fig. 7a) and skewness (see Fig. 7c) there are statistically significant differences for up to 9 and 5 pigs per frame, respectively. Due to the OF vectors pointing in multiple directions and the large number of pigs per frame, the out-of-control pigs' vectors are averaged out, as the number of pigs increases.

3.3. *Comparison of MAHs from moving and stationary pigs*

In this section, we compare MAHs of moving and stationary pigs. We consider two types of out-of-control situations: (1) an animal is stationary or moving slower than it should be moving and (2) an animal is moving too quickly. The number of out-of-control situations can be related to the welfare of the animals. As an illustration, if pigs are encouraged to move too quickly, they will start tripping or stampeding. In contrast, if all animals are moving and some animals are not, this can indicate that animals have some physical problems. Data *set B* (frames with six moving and one stationary pig) is used for comparison of moving and stationary pigs (the data are illustrated in Fig. 8). The results of the comparison are used to detect stationary pigs.

First, using all OF vectors of moving pigs', the median, 25th, and 75th percentiles are estimated. Then, all pigs in data *set B* are identified individually. MAHs are constructed for each identified and fully visible pig.

Figure 9 represents the MAHs of moving and stationary pigs. The moving pigs lie between the 25th and 75th percentiles, while the MAH for the stationary pig does not. If individual pigs can be identified, abnormally moving pigs are detected by comparing individual pigs' MAHs to corresponding angle percentiles. In our case, individual pigs cannot be identified. Thus the abnormal movement is detected by comparing the lengths of individual OF vectors to their corresponding bins.

3.4. *The results of feature selection*

As described earlier, it is difficult to separate pigs when they are close to each other. Therefore, it can be difficult to identify individual animals in a video sequence. Thus, it is not possible to compare histograms from individual pigs in most cases. We suggest analysing individual OF vectors instead. Figure 9 shows that not all bins have the same differentiation power, and since computational power is important in a real-time procedure, we perform feature selection to identify the set of median OF vectors (bins) with the best discrimination power.

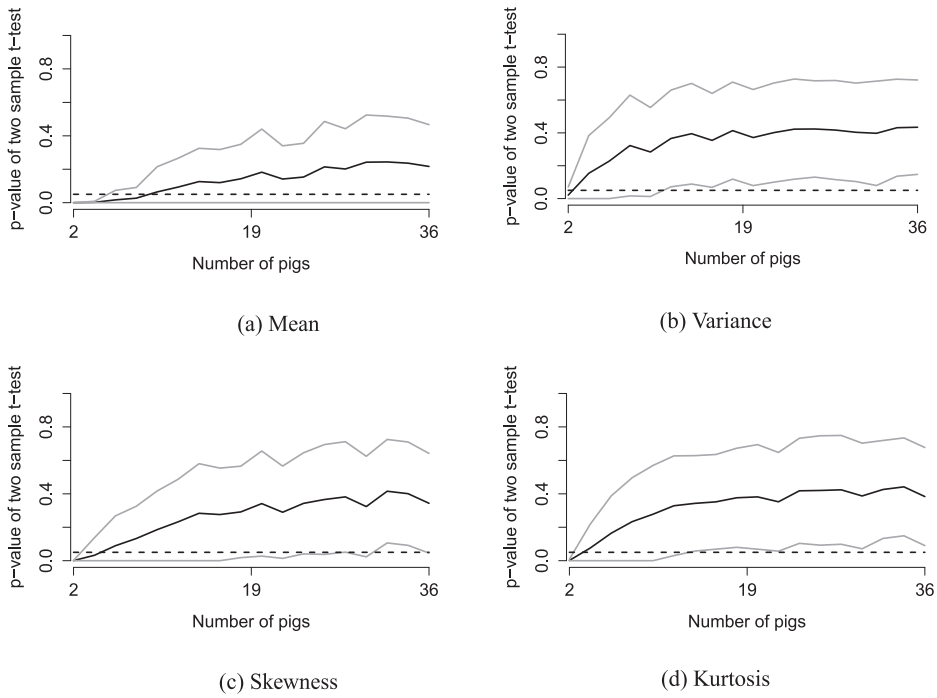


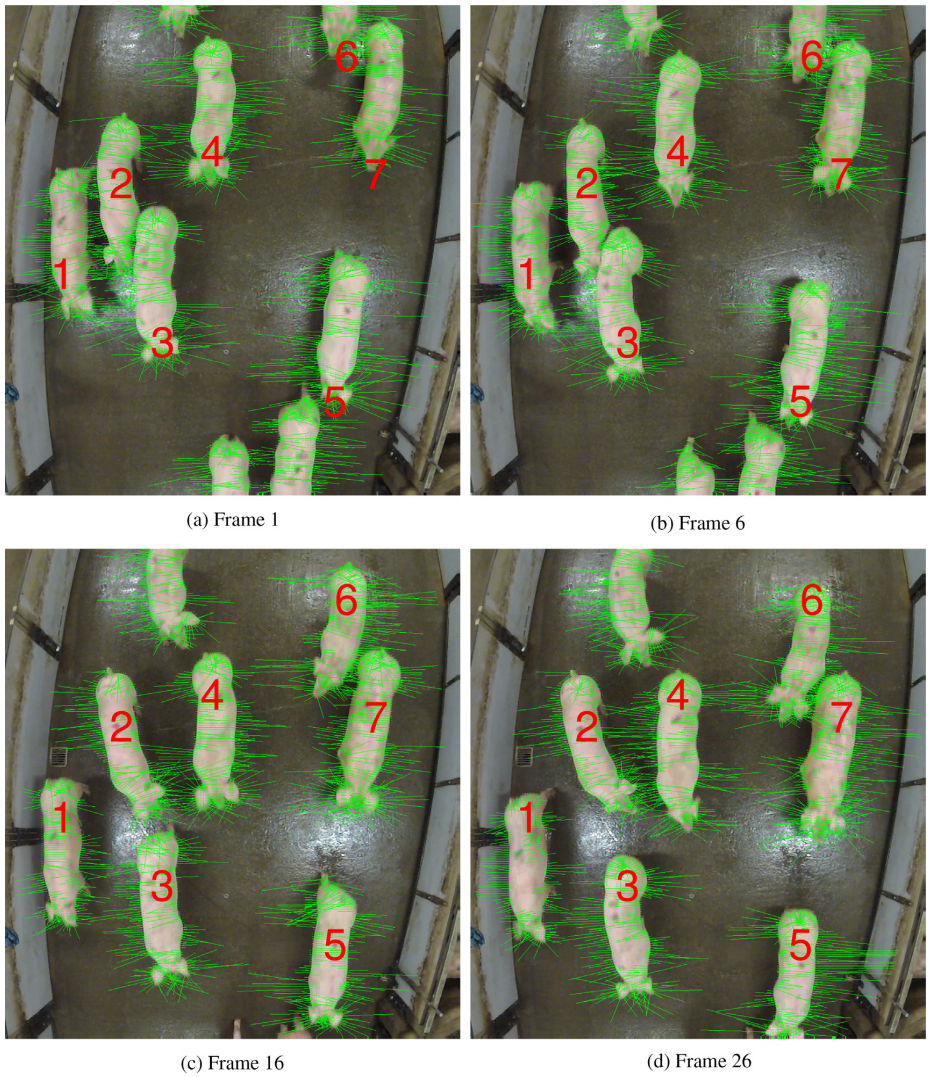
Fig. 7 – *p*-values of the two-sample *t-tests* for the low-level statistics between the two groups. The solid black lines represent mean values while solid grey lines represent ±one standard deviation from the mean. The black dashed line represents 0.05 significance level.

We used the 190 manually annotated frames from *set C* (190 frames of multiple pigs where some of them are stationary) to determine the set of bins for the best two-class separation.

Fully visible pigs are individually detected and annotated in each frame. A total of 540 pigs are annotated, of which 294 pigs are annotated as stationary, whereas 246 pigs are annotated as moving. The sequential selection identifies that the most important bins in the MAH are: 23, 26, 27 and 28 (see Fig. 5). The direction and length of the vectors are shown in Fig. 10. As it can be seen from this figure, all selected bins point in the moving direction.

3.5. *Monitoring and identification of abnormal movement*

In this section, we show how the results of this comparison in Section 3.3 can be applied to locate stationary pigs in a video recording. The same principles can also be applied to identify pigs that are moving at high speed. To identify pigs with abnormal movement, we propose to monitor the four selected angle bins. A vector is declared out-of-control if its length is outside the 25th and 75th percentile range for the corresponding angle. To directly monitor the vectors that are



(a) Frame 1 (b) Frame 6

(c) Frame 16 (d) Frame 26

Fig. 8 – Indexed pigs in the image are used for further analysis. Green lines represent OF vectors. As can be seen from the figures pig number 2 is not moving forward.

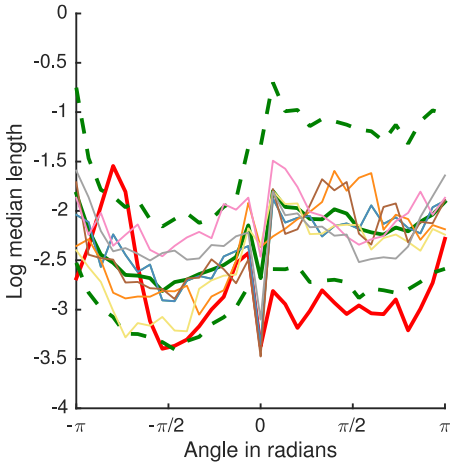


Fig. 9 – MAH of moving pigs: pig 1 – blue, pig 3 – orange, pig 4 – yellow, pig 5 – brown, pig 6 – pink, pig 7 – grey, and stationary pig 2 – red. The median (solid green line), 25th and 75th (green dash lines) percentiles of all moving pigs.

within selected angles using low-level statistics is not optimal, as a few abnormally moving pigs in a large group will be averaged out, just like they would be by applying low-level statistics directly on a vector length. To make the monitoring independent of the number of pigs in a frame, we suggest monitoring the concentration of out-of-control vectors over the frame.

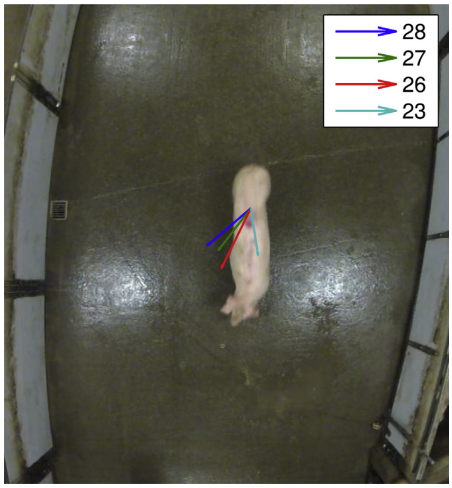


Fig. 10 – The SVM feature selection identified four important angle bins (bin 28, 27, 26, and 23) in MAH, which discriminates moving pigs from stationary pigs. The median vectors of those four angle bins are illustrated in the figure.

It is expected that a certain number of the control pixels will be declared out-of-control in normally moving pigs (false alarms). Thus, we will monitor the density of such occurrences rather than single occurrences. The density of the out-of-control pixels is higher in the areas where abnormal movement is present, in comparison to those areas where normal movement is present. The out-of-control pixel density is tracked within a small patch of 61 × 61 pixels. The density is assigned to the middle pixel of each patch. A 2D interpolation of all patches is used to obtain a smooth density map for each frame.

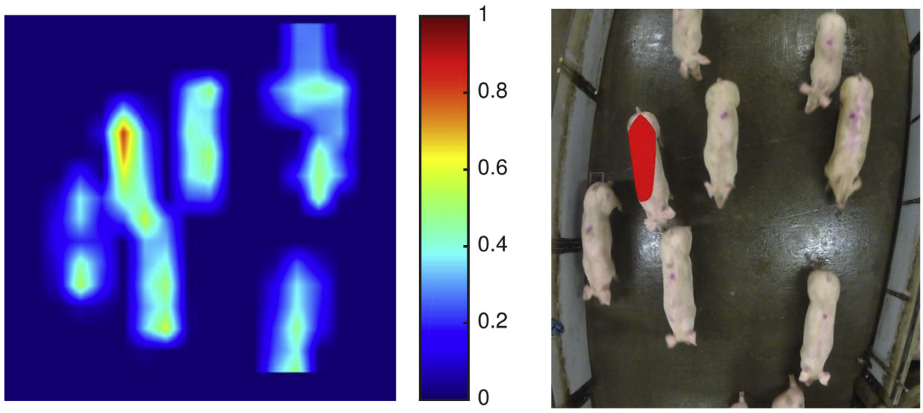
The results of the detection and location of a stationary pig are presented for a video sequence (*set E*). The MAHs of 382 pigs, of which 277 are stationary, and 105 are moving from *set C*, are used for model training. The MAHs of the moving pigs in the *set C* are used to estimate the median, 25th and 75th percentiles. The test *set E* consists of 150 frames, where one pig is stationary in approximately 30 frames. In the experiment, the vectors within the selected angles (corresponding to 23rd, 26th, 27th and 28th bins in MAH) are monitored.

The density map of one of the frames is presented in Fig. 11a. All pigs have some out-of-control pixels. However, one peak covering a large area is higher than the rest. This peak identifies a pig that is not moving. Thresholding is used to separate out the high peak that covers a large area. The threshold value that best separates moving and stationary pigs is determined in a training stage, and set to 0.43. In Fig. 11b, the stationary pig is clearly marked. In 11 frames out of approximately 30, the stationary pig is identified as stationary. In the first 18 frames, the pig is not moving forward, but its body is still moving due to inertia. No false alarms, declaring an out-of-control situation when it is not present, were detected during the experiment. Choosing a lower threshold value would increase the false alarm rate, corresponding to moving pigs being identified as stationary.

3.6. *Performance of modified angular histograms on large animals*

In this section, we illustrate the performance of the proposed monitoring algorithm based on MAH and feature selection. Two groups with different numbers of randomly sampled pigs are estimated using data *set D* (eight frames with 29 moving and two stationary pigs). For each frame, the out-of-control density map is estimated. The thresholding is performed, and the out-of-control areas are determined. The mean, variance, skewness, and kurtosis are estimated for out-of-control areas of each frame, and a two-sample *t-test* is performed for the low-level statistics as in Section 3.2. It is expected that a certain number of pixels will be declared out-of-control over normally moving pigs, but the largest (maximum) area will be over abnormally moving pigs.

The *t-test* indicates that on average we fail to reject the null hypothesis that the samples from the two groups have an equal mean with unequal variances at significance level 0.05 for kurtosis (see Fig. 12c) and also in most cases for the mean (see Fig. 12a). On average the null hypothesis is rejected for the variance statistic (see Fig. 12b) if the number of pigs is less than 20. As discussed earlier, a pig's abnormal behaviour can be averaged out if there is a large number of pigs in the frame.



(a) Density map of out-of-control vectors. (b) Identified stationary pig.

Fig. 11 – Identification of the stationary pig (pig no. 2 in Fig. 8) in the frame. On the left, the density map of the out-of-control vectors is illustrated. The density is highest for the stationary pig on the left figure.

The *t-test* indicates that the skewness and maximum area measures are different between the two groups for all investigated numbers of pigs. An in-depth analysis confirms that the area for the out-of-control pigs will be larger while for the in-control pigs it will be smaller.

4. Discussion

Detecting number of abnormalities in a sequence of frames can be used to enhance animal welfare in two aspects:

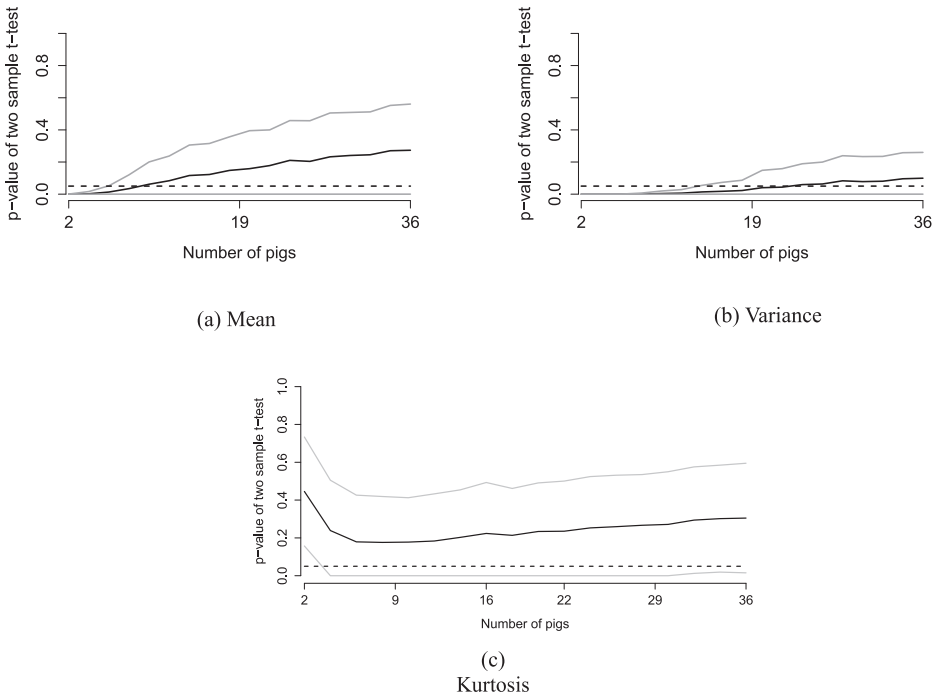


Fig. 12 – *p*-values of the two-sample *t-tests* for the mean, variance and kurtosis of out-of-control area. The solid black lines represent mean values while solid grey lines represent ±one standard deviation from the mean. The black dashed line represents 0.05 significance level.

- To compare the number of abnormalities in herds of pigs that come from different farmers, different drivers who transport the pigs from farm to a slaughterhouse, or during different work shifts. The comparison will identify situations where additional inspection must be performed to determine causes of animal stress. In some recordings, it can be seen that pigs are encouraged to move forward too intensely leading to pigs tripping and stampeding. The cause of rushing can be a lack of training or time pressure. In such cases, an inspection can be performed to identify the reason and action can be taken to improve animal welfare.
- To develop a real-time monitoring that will analyse the number of abnormalities and detect sudden changes in a pig herd. Sudden changes can be a result of an obstacle falling onto a path or some animals that are severely injured or sick. The real-time monitoring would be particularly useful to inform the operators to slow down when encouraging pigs too much.

In SPC, counts of out-of-control situations can be used for cause identification. However, in our case there are two complications. (1) The method should take into consideration whether the estimated location of an out-of-control pig is consistently identified. For example, it is important that approximately the same location is identified as out-of-control in a sequence of frames. The detection of an out-of-control pig in only a single frame is certainly a false alarm. (2) A high correlation between sequential frames in the video is present. Most of the current SPC methods for continuous real-time monitoring have been developed based on the assumption that observations are independent. The frames from videos certainly violate this assumption (unless the frame rate is very low). Thus, there is a need for an SPC method that would take into account correlation between frames. However, this correlation can also be used for making predictions, early detection of events as well as reducing excessive variation due to noise.

Several types of abnormalities are of interest in monitoring pigs in slaughterhouses. When pigs are moving too slowly or not moving when they should be moving can indicate sick or injured animals, while pigs moving too quickly can indicate stressed animals. Our proposed method can locate stationary pigs, but the similar principles can be used to locate pigs moving too quickly or too slow.

5. Conclusion

In this paper, we analyse the movement of individual pigs as well as movement of the entire pig herd, to provide information that can be used for real-time behaviour monitoring and welfare assurance. In addition, the potential differences between monitoring a small vs. large number of animals are discussed. Pig movement is recorded with simple cameras that do not interfere with the usual animal behaviour. Due to similarities in the appearance of the animals, it is computationally expensive to track individual animals. Therefore, the movement of animals in a herd is analysed using optical flow (OF). The analysis indicates that pigs possess a specific walking pattern due to their elongated body shape and trotting movement. Also, a varying number of pigs can be present in a frame and pigs can enter and leave the frame. As a result, low-level statistics of the OF are not efficient for monitoring pigs' behaviour. We suggest using modified angular histograms (MAHs) of the OF for analysis of pigs' movements. The MAHs indicate that a small fraction of all identified OF vectors represents the pig's actual direction and speed.

We propose a method for abnormal movement detection and location in various sizes of pig herds using information obtained from MAHs. We consider abnormalities such as a pig being stationary when it should be moving, as it can indicate an injured animal or an obstacle in the path. The method can also be used for detecting pigs moving either too slowly or too quickly in comparison to normally moving pigs. The results indicate that the method successfully identifies a stationary pig without any false alarms.

Two problems must be solved to make the method work in real-time. The SPC method has to take into account the correlation that is present between frames in a video sequence. This correlation violates the assumption of independence. The consistency in the estimation of out-of-control pigs given their location should also be analysed.

Acknowledgement

This work was carried out as part of the Ph.D. study project Monitoring Animal Wellbeing at the Technical University of Denmark. The project is partly funded by the Danish Meat Research Institute. The authors thank Pia Brandt for collecting and annotating the videos.

REFERENCES

Ahrendt, P., Gregersen, T., & Karstoft, H. (2011). Development of a real-time computer vision system for tracking loose-housed pigs. *Computers and Electronics in Agriculture, 76*, 169–174. http://dx.doi.org/10.1016/j.compag.2011.01.011.

Ali, S., & Shah, M. (2008). Floor fields for tracking in high density crowd scenes. In *Computer vision – ECCV 2008* (pp. 1–14). Springer. http://dx.doi.org/10.1007/978-3-540-88688-4_1.

Andrade, E., Blunsden, S., & Fisher, R. (2006). Modelling crowd scenes for event detection. In *18th international conference on pattern recognition (ICPR)* (vol. 1, pp. 175–178). http://dx.doi.org/10.1109/ICPR.2006.806.

Brandt, P., Rousing, T., Herskin, M., & Aaslyng, M. (2013). Identification of postmortem, indicators of welfare of finishing pigs on the day of slaughter. *Livestock Science, 157*, 535–544. http://dx.doi.org/10.1016/j.livsci.2013.08.020.

Caja, G., Hernandez-Jover, M., Conill, C., Garin, D., Alabern, X., Farriol, B., et al. (2005). Use of ear tags and injectable transponders for the identification and traceability of pigs from birth to the end of the slaughter line. *Journal of animal science, 83*, 2215–2224. doi:/2005.8392215x.

Danish Crown (2014), Departments. URL: http://www.danishcrown.com/page4539.aspx. Accessed 14.08.2015.

Dawkins, M. S., Cain, R., & Roberts, S. J. (2012). Optical flow, flock behaviour and chicken welfare. *Animal Behaviour, 84*, 219–223. http://dx.doi.org/10.1016/j.anbehav.2012.04.036.

Dawkins, M., Lee, H., Waitt, C., & Roberts, S. (2009). Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait. *Applied Animal Behaviour Science, 119*, 203–209. http://dx.doi.org/10.1016/j.applanim.2009.04.009.

Doane, D. P. (1976). Aesthetic frequency classifications. *The American Statistician, 30*, 181–183. http://dx.doi.org/10.1080/00031305.1976.10479172.

Gronskyte, R., Clemmensen, L., Hviid, M., & Kulahci, M. (2015). Pig herd monitoring and undesirable tripping and stepping prevention. *Computers and Electronics in Agriculture, 119*, 51–60. http://dx.doi.org/10.1016/j.compag.2015.09.021.

Guo, Y., Zhu, W., Jiao, P., & Chen, J. (2014). Foreground detection of group-housed pigs based on the combination of mixture of Gaussians using prediction mechanism and threshold segmentation. *Biosystems Engineering, 125*, 98–104. http://dx.doi.org/10.1016/j.biosystemseng.2014.07.002.

Guo, Y.-Z., Zhu, W.-X., Jiao, P.-P., Ma, C.-H., & Yang, J.-J. (2015). Multi-object extraction from topview group-housed pig images based on adaptive partitioning and multilevel thresholding segmentation. *Biosystems Engineering, 135*, 54–60. http://dx.doi.org/10.1016/j.biosystemseng.2015.05.001.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., & Tibshirani, R. (2009). *The elements of statistical learning* (Vol. 2). New York: Springer.

Helbing, D., Johansson, A., & Al-Abideen, H. Z. (2007). Dynamics of crowd disasters: an empirical study. *Physical Review E, 75*, 46–109. http://dx.doi.org/10.1103/PhysRevE.75.046109.

Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. In *SPIE 0281, techniques and applications of image understanding* (pp. 319–331). http://dx.doi.org/10.1117/12.965761.

Huhtala, A., Suhonen, K., Mäkelä, P., Hakojärvi, M., & Ahokas, J. (2007). Evaluation of instrumentation for cow positioning and tracking indoors. *Biosystems Engineering, 96*, 399–405. http://dx.doi.org/10.1016/j.biosystemseng.2006.11.013.

Kashiha, M., Bahr, C., Ott, S., Moons, C. P., Niewold, T. A., Ödberg, F. O., et al. (2013a). Automatic identification of marked pigs in a pen using image pattern recognition. *Computers and Electronics in Agriculture, 93*, 111–120. http://dx.doi.org/10.1016/j.compag.2013.01.013.

Kashiha, M., Pluk, A., Bahr, C., Vranken, E., & Berckmans, D. (2013b). Development of an early warning system for a broiler house using computer vision. *Biosystems Engineering, 116*, 36–45. http://dx.doi.org/10.1016/j.biosystemseng.2013.06.004.

Lind, N. M., Vinther, M., Hemmingsen, R. P., & Hansen, A. K. (2005). Validation of a digital video tracking system for recording pig locomotor behaviour. *Journal of Neuroscience Methods, 143*, 123–132. http://dx.doi.org/10.1016/j.jneumeth.2004.09.019.

Matlab. (2012a). *Computer vision system toolbox(R2012b)*. URL: http://uk.mathworks.com/products/computer-vision/.

Matlab. (2012b). *Image processing toolbox (R2012b)*. URL: http://uk.mathworks.com/products/image/.

McFarlane, N. J., & Schofield, C. P. (1995). Segmentation and tracking of piglets in images. *Machine Vision and Applications, 8*, 187–193. http://dx.doi.org/10.1007/BF01215814.

Montgomery, D. C. (2007). *Introduction to statistical quality control*. John Wiley & Sons.

Nakarmi, A. D., Tang, L., & Xin, H. (2014). Automated tracking and behavior quantification of laying hens using 3d computer vision and radio frequency identification technologies. *Transactions of the ASABE, 57*, 1455. http://dx.doi.org/10.13031/trans.57.10505.

Ng, M. L., Leong, K. S., Hall, D., & Cole, P. H. (2005). A small passive UHF RFID tag for livestock identification. In *Microwave, antenna, propagation and EMC technologies for wireless communications, 2005. MAPE 2005. IEEE international symposium on* (Vol. 1, pp. 67–70). http://dx.doi.org/10.1109/MAPE.2005.1617849.

Oczaka, M., Viazzi, S., Ismayilova, G., Sonoda, L. T., Roulston, N., Fels, M., et al. (2014). Classification of aggressive behaviour in pigs by activity index and multilayer feed forward neural network. *Biosystems Engineering, 119*, 89–97. http://dx.doi.org/10.1016/j.biosystemseng.2014.01.005.

Porto, S. M., Arcidiacono, C., Anguzza, U., & Cascone, G. (2015). The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system. *Biosystems Engineering, 133*, 46–55. http://dx.doi.org/10.1016/j.biosystemseng.2015.02.012.

Porto, S., Arcidiacono, C., Giummarra, A., Anguzza, U., & Cascone, G. (2014). Localisation and identification performances of a real-time location system based on ultra wide band technology for monitoring and tracking dairy cow behaviour in a semi-open free-stall barn. *Computers and Electronics in Agriculture, 108*, 221–229. http://dx.doi.org/10.1016/j.compag.2014.08.001.

Prola, L., Perona, G., Tursi, M., & Mussa, P. P. (2010). Use of injectable transponders for the identification and traceability of pigs. *Italian Journal of Animal Science, 9*, 35. http://dx.doi.org/10.4081/ijas.2010.1341.

Ruiz-Garcia, L., & Lunadei, L. (2011). The role of RFID in agriculture: applications, limitations and challenges. *Computers and Electronics in Agriculture, 79*, 42–50. http://dx.doi.org/10.1016/j.compag.2011.08.010.

Rushen, J., Chapinal, N., & De Passill, A. M. (2012). Automated monitoring of behavioural-based animal welfare indicators. *Animal Welfare, 21*, 339–350. http://dx.doi.org/10.7120/09627286.21.3.339.

Tøgersen, F. A., Skjøth, F., Munksgaard, L., & Højsgaard, S. (2010). Wireless indoor tracking network based on kalman filters with an application to monitoring dairy cows. *Computers and Electronics in Agriculture, 72*, 119–126. http://dx.doi.org/10.1016/j.compag.2010.03.006.

Tu, G. J., Karstoft, H., Pedersen, L. J., & Jørgensen, E. (2013). Foreground detection using loopy belief propagation. *Biosystems Engineering, 116*, 88–96. http://dx.doi.org/10.1016/j.biosystemseng.2013.06.011.

Wedel, D. A., & Cremers, P. D. (2011). *Optical flow estimation*. London: Springer.

Youssef, A., Exadaktylos, V., & Berckmans, D. A. (2015). Towards real-time control of chicken activity in a ventilated chamber. *Biosystems Engineering, 135*, 31–43. http://dx.doi.org/10.1016/j.biosystemseng.2015.04.003.

Appendix C

Appendix 3

R. Gronskyte, L. H. Clemmensen, M. S. Hviid, M. Kulahci, "Pig herd monitoring and undesirable tripping and stepping prevention", *Computers and Electronics in Agriculture*, vol. 119, pp. 51-60, 2016.

Computers and Electronics in Agriculture 119 (2015) 51–60

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Pig herd monitoring and undesirable tripping and stepping prevention

Ruta Gronskyte [a,*], Line Harder Clemmensen [a], Marchen Sonja Hviid [b], Murat Kulahci [a,c]

[a] DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark
[b] Danish Meat Research Institute, Taastrup, Denmark
[c] Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

ARTICLE INFO

Article history:
Received 17 April 2015
Received in revised form 20 July 2015
Accepted 24 September 2015
Available online 28 October 2015

Keywords:
Abnormal behavior detection
Animal welfare
Optical flow
Pig unloading
Video analysis

ABSTRACT

Humane handling and slaughter of livestock are of major concern in modern societies. Monitoring animal wellbeing in slaughterhouses is critical in preventing unnecessary stress and physical damage to livestock, which can also affect the meat quality. The goal of this study is to monitor pig herds at the slaughterhouse and identify undesirable events such as pigs tripping or stepping on each other. In this paper, we monitor pig behavior in color videos recorded during unloading from transportation trucks. We monitor the movement of a pig herd where the pigs enter and leave a surveyed area. The method is based on optical flow, which is not well explored for monitoring all types of animals, but is the method of choice for human crowd monitoring. We recommend using modified angular histograms to summarize the optical flow vectors. We show that the classification rate based on support vector machines is 93% of all frames. The sensitivity of the model is 93.5% with 90% specificity and 6.5% false alarm rate. The radial lens distortion and camera position required for convenient surveillance make the recordings highly distorted. Therefore, we also propose a new approach to correct lens and foreshortening distortions by using moving reference points. The method can be applied real-time during the actual unloading operations of pigs. In addition, we present a method for identification of the causes leading to undesirable events, which currently only runs off-line. The comparative analysis of three drivers, which performed the unloading of the pigs from the trucks in the available datasets, indicates that the drivers perform significantly differently. Driver 1 has 2.95 times higher odds to have pigs tripping and stepping on each other than the two others, and Driver 2 has 1.11 times higher odds than Driver 3.

1. Introduction

Today most consumers are not interested in meat quality alone but also in the welfare of the animal (Verbeke and Viaene, 2000; Napolitano et al., 2010; Kehlbacher et al., 2012). Moreover, high levels of stress before slaughter can affect the meat quality (Warriss et al., 1994; Brandt et al., 2013). Bruises caused before slaughter can also impair the value of meat sold with skin. In this paper, the behavior of pigs in a large and automated slaughter house that can handle up to 62,000 pigs per week is analyzed. It is difficult to keep track of that many pigs and ensure that they remain stress-free during handling. To complicate matters further, pigs which are unfamiliar with each other (e.g. coming from different fattening) can potentially attack each other causing unnecessary stress and physical damage (Oczak et al., 2014). In this study, pigs are filmed upon arrival at the slaughter house during the unloading of a truck. During transportation, pigs are usually sedentary therefore they often move slowly during unloading. To speed up the process, truck drivers are allowed to use specially designed sticks with sound effects. Pigs often react differently and some start moving too fast, resulting in a stampede. Consequently, they can start tripping and may step on each other. This situation can increase the stress level and may cause injuries (Broom, 2005). Most of the undesired situations, such as pigs tripping or stepping on each other, happen when animals are moving too fast and densely confined in an area. Consequently, we want a process where they move as fast as possible while ensuring that the animals remain stress-free and unharmed. The aim of this paper is to monitor pig herds at the slaughterhouse and determine when pigs stampede or about to stampede and give notice to the personnel to slow down the unloading and avoid events where pigs are tripping or stepping on each other. In our previous work, we proposed monitoring pig herd movement based on video surveillance without identifying individual animals (Gronskyte et al., 2015). Previous studies by Ahrendt et al. (2011) and Kashiha et al. (2013) tracked several individual pigs in a constrained area,

* Corresponding author. Tel.: +45 45 25 53 51.
E-mail address: rgro@dtu.dk (R. Gronskyte).
URL: http://www.dtu.dk (R. Gronskyte).

http://dx.doi.org/10.1016/j.compag.2015.09.021
0168-1699/

which means that all pigs at all time were visible, whereas we analyzed the herd movement in an "unconstrained" area, allowing for pigs to enter and then leave the surveyed area.

The videos provided for this analysis were recorded at an angle with radial lens distortion. As a result of lens and foreshortening distortions, some additional steps had to be taken before the final behavior could be classified. Optical Flow (OF) was used to monitor the pig herd movement. Initially OF is estimated for an entire frame, thus pigs are identified and OF is filtered to only analyze movement of the pigs. Due to the special trot of pigs, averages of OF vectors are not suitable for behavior monitoring (Gronskyte et al., 2015) as proposed by Dawkins et al. (2009, 2012). Instead, we propose in Gronskyte et al. (2015) monitoring discriminant herd movement features extracted from a modified angular histogram (MAH) (Gronskyte et al., 2015) using Support Vector Machines (SVM) (Hastie et al., 2009, pp. 417–458). A MAH of a single frame describes the distribution of the OF vectors' median lengths over the entire angle range. The analysis based on MAH can be used in cases where monitored objects present a specific walking pattern. The overview of the described approach is presented as a flowchart in Fig. 1. The results of the SVM can also be studied on the basis of several frames, rather than a single frame. In this context, we will describe the possibilities of identifying the causes of the stampede.

In the following section a brief literature review on behavior monitoring is presented. In Section 2 the data are presented. The method is presented in Section 3 followed by the summary of the results given in Section 5 and the conclusion in Section 6.

1.1. Literature review

There is extensive research carried out on behavior monitoring of humans, and farm, wild and laboratory animals. There have primarily been two approaches to human behavior monitoring: tracking individuals or using OF. An overview of the human behavior monitoring methods is given in Hu et al. (2004). Some of the methods focus on detecting a single action, like a snatch, in surveillance videos (Ibrahim et al., 2010) or dense crowds trying to use escalators (Ihaddadene and Djeraba, 2008). A large number of methods focus on detecting any abnormality in the videos (Mehran et al., 2009; Kratz and Nishino, 2009; Boiman and Irani, 2007). Animal identification is an important step in most animal monitoring methods, and it has been implemented in a variety of different ways for different animals. Fish and rats are usually monitored (Spink et al., 2001; Ardekani et al., 2013) in a highly controlled environment, possibly using additional markers to identify individual animals, with the primary focus on monitoring repetition of a specific action and/or the duration of the action. Some research has also been done in monitoring insects, but they are particularly challenging subjects to track due to their small size and similarity (Hendriks et al., 2012). In bat movement analysis (Breslav et al., 2012) the main interest is to identify individual animals and compare their flying trajectories.

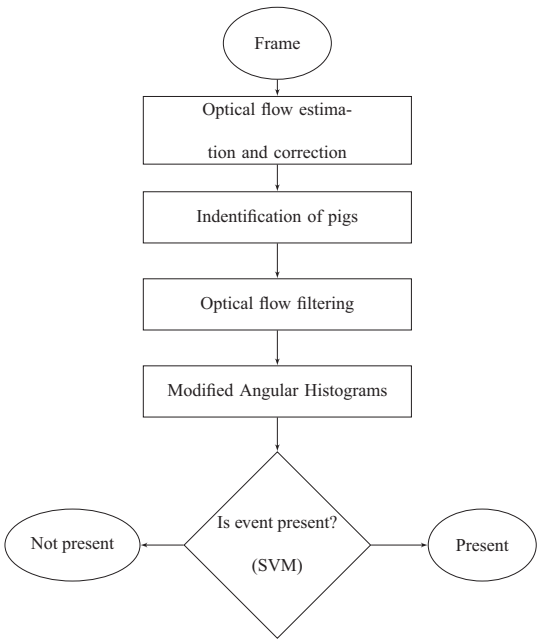


Fig. 1. Flowchart of the proposed pig behavior monitoring approach.

More research has been done in monitoring cows', chicken broilers' and pigs' behavior. A study by Cangar et al. (2008) used a model-based monitoring tool to track locomotion and posture of pregnant cows' in a pen. Studies by Dawkins et al. (2009, 2012) and Roberts et al. (2012) on chicken behavior analysis use low level features of OF to monitor chickens' welfare. Pigs can be tracked using radio tags that are attached to the animal's ear (Ng et al., 2005) or injected in the animal's body (Prola et al., 2010; Caja et al., 2005), or using video surveillance. For example, Tu et al. (2013) and Guo et al. (2014) use background subtraction methods to identify pigs in a frame. Their methods can handle illumination changes, which is different from our case that involves controlled illumination. Pigs were also tracked in a pen using markers (Kashiha et al., 2013), features (Ahrendt et al., 2011) or shape matching (Tillett et al., 1997). The feature based approach successfully tracks pigs without loosing them for eight-minutes. The computer vision methods are used to monitor not only the pig' behavior but also the activity levels in relationship to the climate in a pen (Costa et al., 2014), to measure pigs' weight (Kashiha et al., 2014) and to analyze pigs locomotion (Kongsro, 2013). None of the methods for animal monitoring in the literature analyze cases where the animals can enter and leave the surveyed area. Such unconstrained areas are on the other hand very common in human crowd monitoring (Kratz and Nishino, 2009; Perko et al., 2013) and the corresponding analysis is often carried out using OF as we also propose in our study.

2. Animals and setup

The pigs were recorded at a Danish slaughterhouse during the unloading process from the truck. The pigs were transported for around 2 h from the fattening houses to the slaughterhouse in commercial three-deck trucks. The pigs in the truck were divided into pens of 15–23 pigs each. There were 9–15 adjustable size pens in a truck. The majority of the pigs are crossbreeds between Duroc, Danish Landrace and Yorkshire (Dx(LxY)). The recordings mainly contain white-skinned pigs with a few colored pigs. Pigs weighed 100–110 kg and were approximately 6 months of age.

The setup of the unloading area is shown in Fig. 2. The pigs are transferred from the trucks to unloading dock and from there they enter the slaughterhouse. The average temperature of 14.8 °C was recorded during the unloading.

A GoPro HERO2 (©2013 Woodman Labs, Inc) camera was used for recording. A sample frame of the videos recorded is given in Fig. 3. The entire frame is not analyzed as the pigs can only move in a certain area of the frame.

Three videos, each of length 17 min 35 s, were recorded at the rate of 29.97 frames per second. The frame height is 1920 and

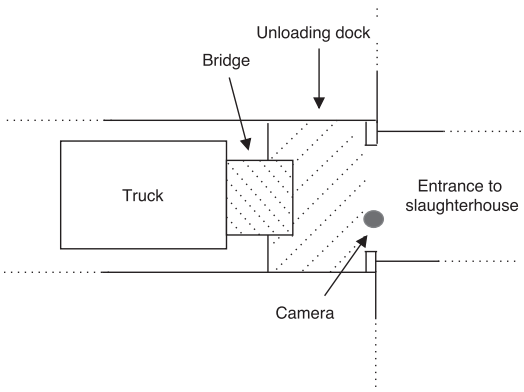


Fig. 2. The setup scheme.

Fig. 3. The sample frame of the recorded videos. The red dash line represents an area of interested and pigs moving direction is considered to be between two arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

width 1080, but the frames shown in this paper are cropped for proprietary reasons. Every frame is analyzed and optical flow is estimated using 5 frame reference delay. In each video, different truck drivers were unloading the pigs.

3. Methods

3.1. Data annotation

Some of the videos contain recordings of a preparation for the unloading which are not analyzed. In total 30,124 frames, where the unloading operation is taking place, are analyzed and annotated. During the unloading, pigs move from the truck towards the entrance of the slaughterhouse as depicted in Fig. 3 and in this paper this is referred to as pigs' movement. Each frame is classified by two classifiers through manual annotation. The first classifier describes the movement of the pigs with three classes: 0 when there is no undesirable event present, 1 when pigs are tripping and 2 when pigs are stepping on each other. The second classifier describes the presence of the truck driver: class 0 when the truck driver is not present in the scene, class 1 when truck driver is present but he is not encouraging the pigs and class 2 when the truck driver is present and encouraging the pigs. The main focus is on the three sub-classes: (1) when the driver is not present and pigs are moving without any undesirable events (3718 frames), (2) when the driver is encouraging the pigs to move causing them to trip (541 frames), and (3) when the driver is encouraging the pigs to move and they are stepping on each other (824 frames). This classification can be used during the unloading and alert the staff when an undesirable event is happening or about to happen. In most of the analyzed cases, the undesirable event is a result of stampede. For the rest of the paper we use the term "event" to represent an undesirable event, such as pigs tripping or stepping on each other. Each event is presented in a series of frames and the number of frames depends on the length of the event.

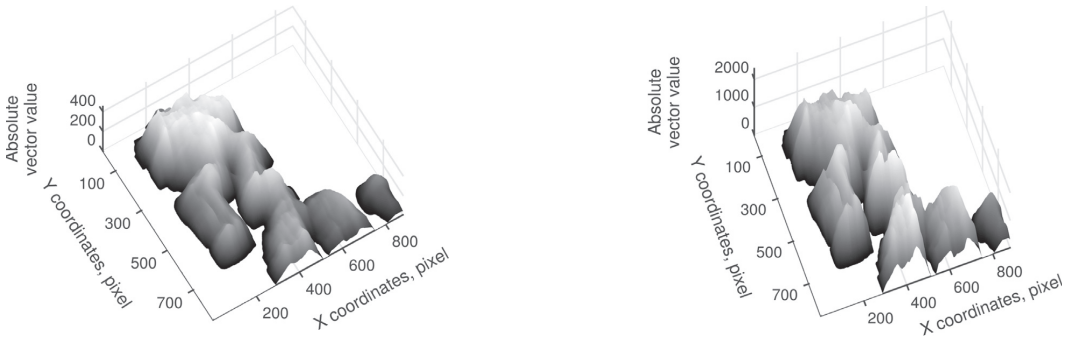
3.2. Movement analysis

In this section, the proposed pig movement analysis is presented. This analysis consists of three steps: (1) motion estimation, (2) feature extraction and (3) frame classification. To estimate the motion of the pigs, we use four steps: (i) optical flow estimation, (ii) identification of pigs, (iii) optical flow filtering, and (iv) distortion correction.

3.2.1. Optical flow determination and correction

In the analysis, MATLAB's Computer Vision System toolbox MATLAB (2015) which employs the Horn–Schunck method with a five frame reference delay is used. As mentioned earlier, the OF vector values had to be adjusted due to radial lens and foreshortening distortions. The former, can be split into two types: (a) pincushion distortion and (b) barrel distortion, which is present in the videos in this study. The barrel distortion maps an image around a sphere, thus a pig's magnification decreases with distance from the optical axis. The latter is a result of perspective projection, thus a pig exiting the truck looks smaller than it exits the area of interest. Fig. 4a shows that the absolute values of the OF vectors are shorter at the lower part of the image than at the upper part. The values in Fig. 4 are smoothed using a finite impulse response filter.

Only a few correction algorithms have been developed because it would often be easier to adjust the camera or to account for the distortion using lens specification. However, since this could not be done for the available data, we needed to adjust the data for lens and foreshortening distortions. Previous work by Altunbasak et al. (2003) suggests to combine lens distortion correction and OF determination. To account for both distortions, the pig length and width are used as reference measurements because of the similarity of the pigs in shape and size. Distortions in both X and Y directions for each pixel in the frame are estimated using length and width of 101 forward moving pigs respectively. This is done



(a) Before. (b) After.

Fig. 4. OF vectors absolute values before and after OF field correction.

because a solid reference is not available. The pigs used for reference are selected according to the following criteria: (1) the pig has to stand straight, so that there is a straight line between the tip of the nose and the tail, (2) the four points of interest (the tip of the nose, the tail and the shoulders) are not occluded and clearly visible, and (3) the line connecting the pig's tail and nose is parallel to the *Y*-axis and the line connecting the shoulders is parallel to the *X*-axis. The length of the pig is calculated from the tip of the nose to the end of the tail of the animal and is assigned to the pixel corresponding to the tip of the nose pixel whereas the width is calculated as the width of the shoulders of the animal and assigned to the furthest left pixel of the shoulder. A polynomial model is used to model the changes in *X* and *Y* directions. The smallest sum of squared errors is achieved using a 3rd order polynomial given as:

$$Z(x,y) = \sum_{j=1}^{4}\sum_{i=1}^{4} p_{ji} x^{4-j} y^{4-i} \quad (1)$$

where x and y are the vector positions in the frame, Z is the length or width of the pig and p_{ji} are polynomial coefficients. Altunbasak et al. (2003) also uses a 3rd order polynomial to correct for lens distortion. Surface plots for distortion in the *X* and *Y* directions can be seen in Fig. 5.

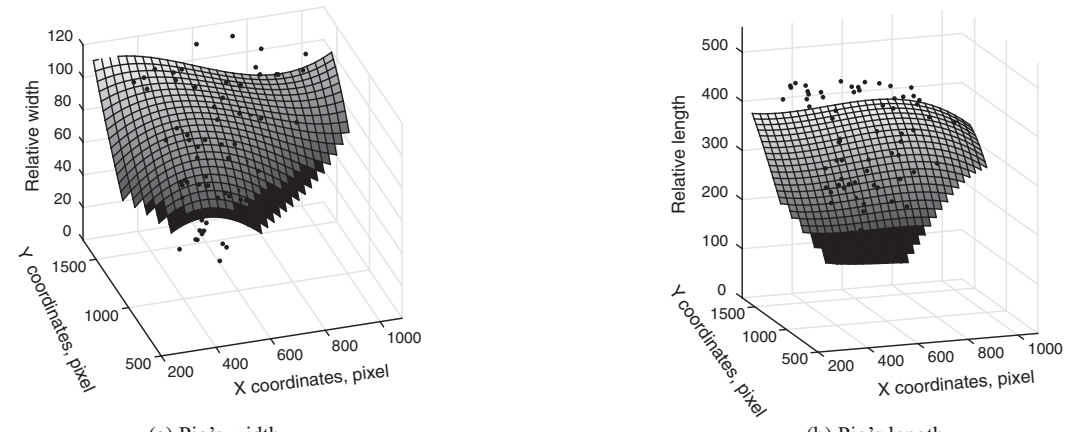
The length and width of each pixel in the region of interest are estimated using Eq. (1) and shown in Fig. 6 together with the identified focal point. The focal point is where the predicted length and width are shortest. The predicted measures are then used to estimate the correctional coefficient field. The function $f(x)$ re-scales the $I = \{X, Y\}$ into interval $[a; b]$ as follows:

$$f(x) = \frac{(b-a)(x - min(I))}{max(I) - min(I)} + a \quad (2)$$

where $a = 1$ because it is expected that the OF vectors in the area with no distortion would remain unchanged and vectors in areas with distortion would be scaled to $b = \frac{max(I)}{min(I)}$. The distortion correction field is used to adjust the OF vectors. This is done by multiplying real values of the OF vector with the width distortion correction coefficients and the imaginary value with the length coefficients. Fig. 4b shows the absolute values of the OF vectors after the correction.

3.2.2. Optical flow filtering

To monitor the pigs' movements, the OF vectors representing movement of the pigs need to be separated from the background. At first, thresholding the pixel color values was considered.



(a) Pig's width. (b) Pig's length.

Fig. 5. The length and width of the pigs over the region of interest (marked as dots), and the 3rd order model fit used for correction of lens and foreshortening distortions.

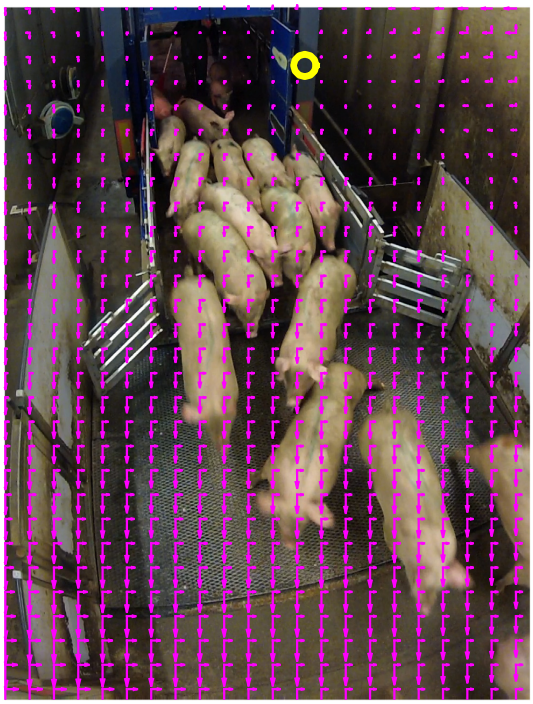


Fig. 6. Lens and foreshortening distortions in the full frame are marked with magenta arrows. The yellow circle represents lens focal point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

However, this method is not feasible due to light reflections, as the OF vectors of the light reflection or light source can have the same length as the pigs' movements. The approach employed to identify individual pigs includes image color map adjustment, color filtering, blob detection, image dilation and holes filling. The image color map adjustment (MATLAB, 2015) increases the contrast in an image, thus increasing the ability to identify the pigs. To completely separate the pigs from the background, blob detection is used. Blob detection (Bailey, 2011; Tuytelaars and Mikolajczyk, 2008) identifies regions in an image that differ from the surrounding areas. In addition, the region on which the pigs can walk is selected out of the entire frame. As a result, all background is eliminated. Image dilation (Soille, 2003, p. 68), which utilizes square structural element with a 5 pixel distance from the center to the edge, is used to include all pixels at the edge of the pigs' bodies. Holes filling (Soille, 2003, p. 208) is used to fill the holes on the pigs' bodies that appear due to markings or dirt.

3.3. Herd movement features

Due to elongated body shape and movement of the pigs, a lot of OF vectors do not represent pigs' actual movement direction and speed. A pigs' body movement resembles the letter "S", therefore the angles of the OF vectors are in all directions (Kongsro, 2013). The vectors corresponding to the actual movement direction and speed are represented in a valley of the MAH. Therefore we suggest to use MAH to analyze pigs' movement. Angle range is divided into bins from $-\pi$ to π using Doane's formula (Doane, 1976) for non-normal distribution:

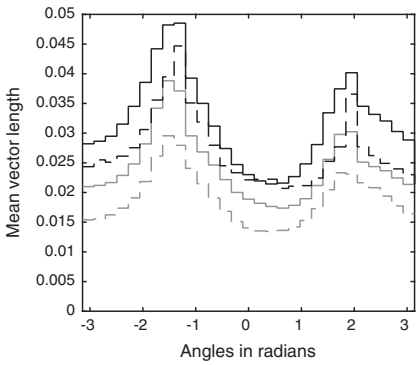
$$k = 1 + \log_2(n) + \log_2\left(\frac{|g_1|}{\sigma_{g_1}}\right) \tag{3}$$

$$\sigma_{g_1} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} \tag{4}$$

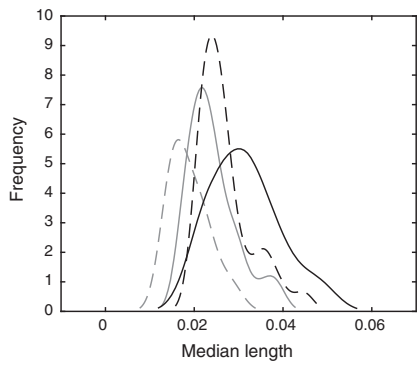
where g_1 is the sample skewness, n is the number of observations, k is the number of bins. These two values are taken as the averages of skewness and the number of OF vectors in the training set respectively. In each bin, the average length of the vectors is registered. Fig. 7a shows the four examples of mean MAH for each annotated sub-class described in Section 2 and Fig. 7b shows the empirical distributions of mean MAH for the four sub-classes.

Fig. 7a and b indicates that when pigs are moving without any events they tend to move slower and when an event is present pigs on average tend to move faster. It is clear that peaks of freely moving pigs are lower than the peaks when an event is present. The difference between the peaks and the valleys of the histograms changes within a sub-class.

The main focus is to estimate the relative measure of density and not necessarily the number of pigs. The pig density is estimated by counting the number of pixels that are classified as pigs.



(a) Mean MAH for each sub-class.



(b) Frequencies of mean MAHs.

Fig. 7. (a) Mean MAH of four annotated sub-classes, and (b) empirical frequency histograms of sub-class MAH. The dashed gray line indicates freely moving pigs, solid gray line is for encouraged pigs but moving without events, dashed black line is for encouraged pigs that are tripping and solid black line is for encouraged pigs that are stepping on each other.

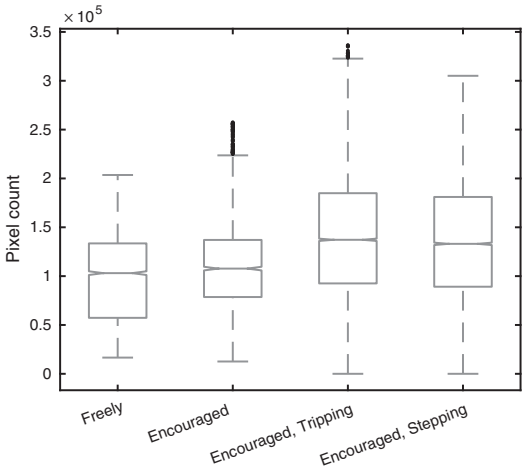


Fig. 8. Pixel counts for different pig behaviors.

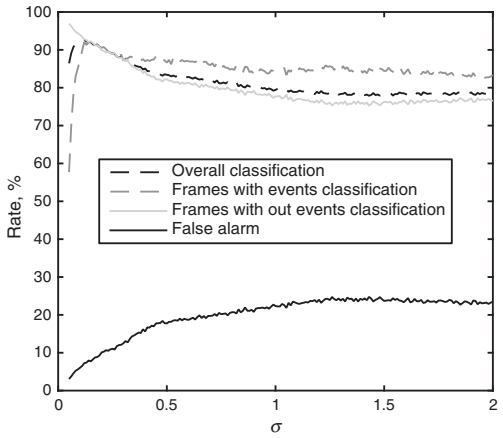


Fig. 10. Identification of optimal scaling parameter σ for SVM.

Fig. 8 shows that when the pigs are moving freely there are usually fewer pixels of pigs than when they are being encouraged to move.

For further analysis, we consider the following four variables: the average of the two highest bins (1) in the left peak and (2) in the right peak, (3) the average of two smallest bins in the valley and (4) the number of pixels that are classified as pigs.

3.4. Movement classification

Figs. 8 and 9 show the differences in the average values of the selected variables between sub-classes. However, as it can be seen in Fig. 9, none of the variables separate the freely walking pigs clearly from the pigs affected by events. We suggest to combine information from all four variables to obtain a clear separation among the sub-classes. For further analysis, the two sub-classes containing events are merged to obtain the identification when stampede is happening or about to happen and adequate actions should be taken.

SVM with radial kernel basis function is applied to separate the different sub-classes. To identify the optimal scaling parameter σ, we randomly divide three selected sub-classes into training (75% of the frames) and testing (25% of the frames) datasets and for each σ, SVM model is estimated and error rates are recorded. The average error rates are shown in Fig. 10. The optimal σ is calculated to be 0.11 and the method's overall classification rate is 93%, while 90% of the frames with events are classified correctly and 94% of the frames without events are classified correctly.

4. Identification of assignable causes

Systems and processes are generally subjected to two types of variation: common cause variation and assignable cause variation. The former is defined as the variation that occurs due to chance and it is perpetually present in any process. The latter, however is due to unusual and unexpected disturbances, also known as assignable causes which are to be detected and removed in order to bring the process back to normal operating conditions. In our study, we assume that a certain amount of stress is likely to be imparted to the animals, irrespective of the handling conditions. However, additional stress that can for example cause a stampede can be introduced due to varying behaviors of truck drivers, differences in farms from which the pigs are coming from, problems with the bridge connecting the truck to the unloading area, etc. It will therefore be of great value to identify the presence of these assignable causes in order to take actions, to alleviate their impact. In industrial processes this is achieved by employing statistical process control (SPC) techniques designed to monitor (real-time) the performance of the processes and declare an out-of-control signal (e.g. a stampede is happening or about to happen) when more

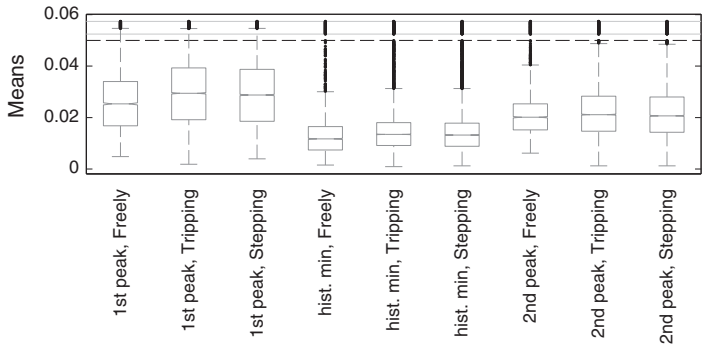


Fig. 9. Selected variables for monitoring animal movement.

than an expected amount of variation is detected. A more in depth implementation of SPC techniques in our study constitutes our future work. In this section we provide a prelude to the identification of assignable causes with an off-line approach and leave the real-time implementation of appropriate statistical process control techniques for our upcoming work.

In our case, our primary concern is the number of out-of-control frames, i.e. frames where undesirable events such as tripping and stepping on each other are present. The methodology presented in Section 3.2 provides classification of each frame. A measure of interest can then be the ratio of out-of-control frames to the total number of frames. Under the assumptions that the probability p of a frame being classified as out-of-control is the same for all frames and that the frames are independent, the number of out-of-control frames in a sample of n frames is binomially distributed as

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{5}$$

where x is a number of frames classified as out-of-control.

The drivers use specially designed sticks to encourage pigs to move faster, if the pigs are heavily encouraged, it can cause stampede. Hence to investigate the influence of the drivers on the behavior of the pigs for example, we divide our data set in three subsets according to drivers that are unloading the trucks. Each driver is classified using the method described in Section 3.2 and the number of out-of-control frames is counted. In total, 16632, 3558, 9934 frames are analyzed for Drivers 1, 2 and 3 respectively. We then build a logistic regression model between dependent (ratios of out-of-control frames and total number of frames per driver) and independent variable (drivers). The estimated coefficients are presented in Fig. 11.

Pairwise comparisons of the drivers' influence are performed using the analysis of variance (ANOVA) (Johnson et al., 2010) as the method takes into account the different sample sizes, and the results are presented in Table 1. The p-values indicate that the drivers are all different at a 5% significance level, even though Fig. 11 shows that Drivers 2 and 3 are more similar than Driver 1. In depth analysis suggests, that Driver 1 has 2.79 and 3.10 times higher odds to have an out-of-control frame (i.e. an event) than Drivers 2 and 3, respectively. Driver 2 has only 1.11 time higher odds than Driver 3.

As we stated earlier, this analysis is performed off-line and therefore is of limited use. On the other hand it provides useful information in terms of identifying the causes for excessive variation that can be easily reduced to normal levels by for example providing better training for the drivers. We should also keep in mind that the binomial distribution assumes independent observations. Clearly assuming successive frames to be independent is not realistic. We are currently in the process of developing better

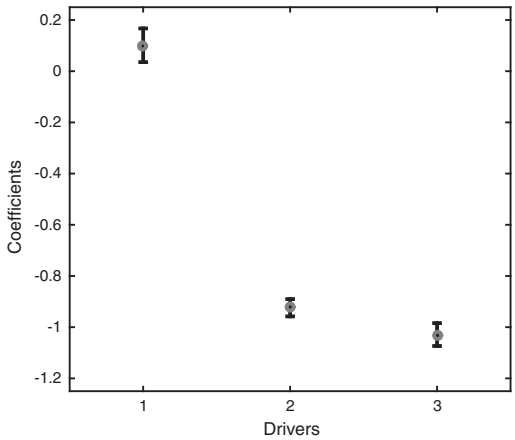


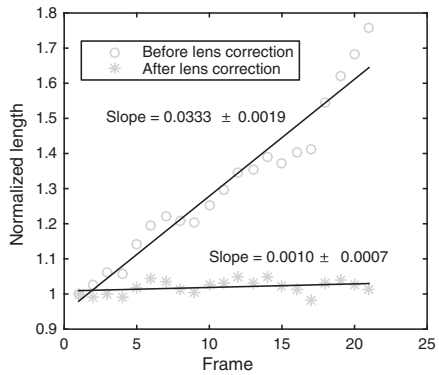
Fig. 11. Estimated coefficients and their confidence intervals for the logistic regression model.

Table 1
The p-values for the ANOVA test comparing the drivers.

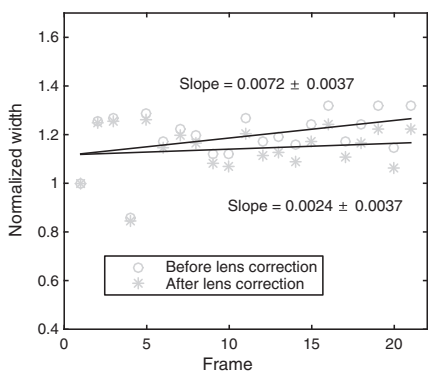
	p-value
Driver 1 vs. 2	<0.001
Driver 1 vs. 3	<0.001
Driver 2 vs. 3	0.00024

Table 2
Classification performance comparison before and after corrections.

	Before (%)	After (%)	p-value
Total classification rate	87	92	<0.001
Frames with events classification rate	67	90	<0.001
Frames without events classification rate	94	94	0.75

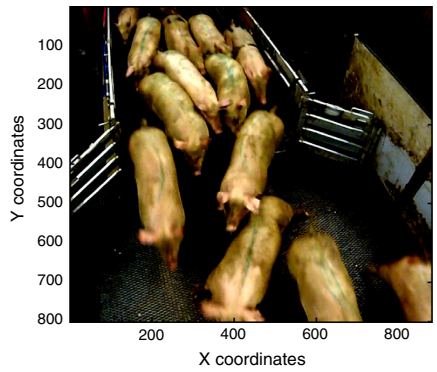


(a) Pig'slength.

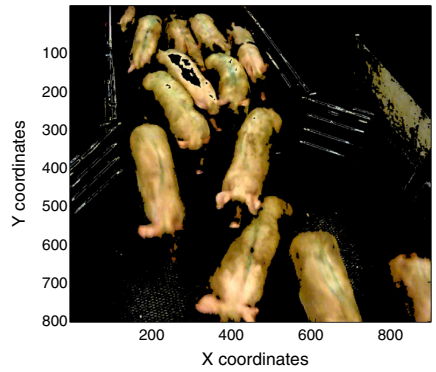


(b) Pig'swidth.

Fig. 12. The length and width of one pig sampled in 21 frames. In frame 1 the pig is fully visible in the area of interest for the first time and in frame 21 for the last time.



(a) Color adjustment.



(b) Color thresholding.

Fig. 13. Two steps in pig identification: (a) color adjustment and (b) pixel color valued thresholding. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

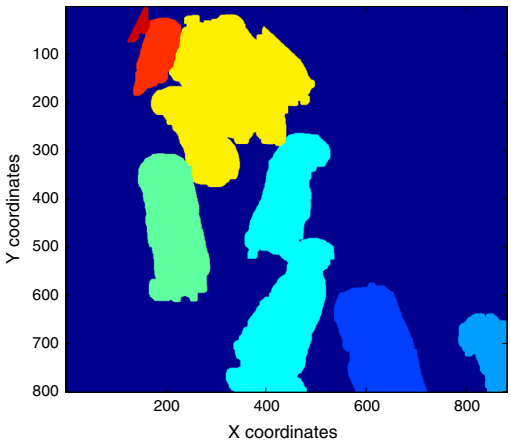


Fig. 14. Final result of pigs identification in an image. Blue color represents background, and other colors represent the pigs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

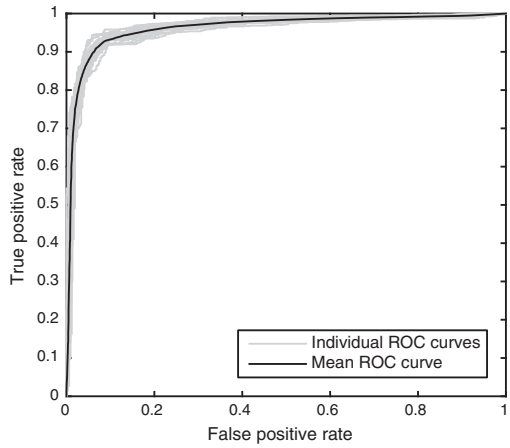


Fig. 15. ROC curves for SVM classification model. The gray lines indicate the performance of the individual randomize test curves, while the black line represents the mean curve.

methods performing real-time monitoring of this type of data. Nonetheless, this rough analysis is provided as a starting point to raise awareness about the techniques available to researchers in monitoring system performance through image data.

5. Results

In this section the results of lens correction, pig identification and movement classification are presented.

5.1. Lens correction

In Section 3.2.1, the method for foreshortening and lens distortion correction is presented. In this section, the effectiveness of the correction is discussed. As it can be seen in Fig. 4b, the high absolute values of OF vectors are distributed over the entire image. Before the correction the high values were concentrated in the top part of Fig. 4a. The first indicator of a good correction is the location of the identified focal point Fig. 6. Two indirect tests were carried out to evaluate the performance of the correction. First the length and width of a randomly selected single pig are sampled every 4 frames spanning from the pig entering to exiting the area of interest. In total 21 width and length measurements are taken. The normalized measurements before and after the correction are shown in Fig. 12. Before the correction the length linearly increases as the pig moves over the bridge. After the correction the length becomes stable. The width of the pig is almost constant both before and after the correction.

The second test compares the performance of event detection in the recordings with and without correction. Three performance measures are compared: (1) total classification rate, classification rates of (2) frames with and (3) frames without events. Both the corrected and uncorrected recordings were randomly divided into training and testing sets. Table 2 compares the classification rates with and without correction using the two-sided t-test. The p-values indicate a significant increase in the total classification rate and classification rate of frames with events after correction. The classification rate for frames without events is not affected by distortions.

5.2. Optical flow filtering

As it can be seen in Fig. 6, most of the pigs are dirty, therefore it is not easy to separate them from the dark background. Adjusting the intensity of the image to between [0.3; 0.8] enhances the pigs in the image. Results of the adjusted color intensity can be seen in Fig. 13a. Using pixel color value thresholding, most of the background was removed in the image, as shown in Fig. 13b. Some errors appeared on the rails of the bridge, on some pigs' backs where dirt and markings are present, at the edges of the pigs' bodies, and at the noses of the pigs. The final result can be seen in Fig. 14.

5.3. Movement classification

In this section movement classification results are presented. The overall classification rates is $92.47 \pm 1.04\%$, while the model sensitivity is $93.5 \pm 1.12\%$, specificity is $89.72 \pm 2.68\%$ and false alarm rate is $6.5 \pm 1.12\%$. The area under receiver operating characteristic (ROC) curves in Fig. 15 is $96.22 \pm 0.52\%$. All the performance measurements indicate a good classification model. In addition, 12 separate events were present in the analyzed videos and the proposed method identified some of the frames of all events. There are several reasons for the method not identifying the rest of the frames:

- There is no clear cut indication for when an action is over and another starts, thus the transition between actions is difficult to classify. For example, when an animal stops walking, its body may still be moving due to large amounts of fat and also the body shape.
- A stationary pig can still move its head and this can be registered as a "movement".

6. Conclusion

The main goal of this study is to prevent undesirable events that can affect animal wellbeing in a slaughterhouse. Events such as pigs tripping or stepping on each other are of the main interest, and they are usually a result of stampede. The first hurdle is to convert image data into numerical data that can be used for classification of each frame based on the behavior of pig herd during the unloading process. In previous works, pigs were individually identified and tracked. This method is not appropriate as there are too many pigs in the surveyed area and pigs can leave and re-enter the area. The chosen method for motion extraction from the images is the optical flow (OF). To quantify the pigs' movement we use modified angular histogram (MAH). We show that MAH is very useful in understanding pigs' body movement and analyzing general movement direction and speed. We expect that MAH can be used in any behavior monitoring case, where surveyed objects present a specific walking pattern. Support Vector Machines (SVM) are applied to selected angle histogram features and to the relative number of pigs in the frame to classify pigs' movements in each frame. We also propose that the outcome of SVM can be used for real-time monitoring. Once the classification is done, we then test each data set selected according to different drivers.

Due to the quality of the recordings we had to consider a problem that was not commonly discussed in the literature: correction of lens and foreshortening distortions using moving reference objects. The proposed algorithm is fast, easy and has proven to be effective.

For future work we are considering to investigate the number of successive out-of-control frames needed to make a decision to alert the staff for real-time herd monitoring. In addition, the analysis of the assignable causes will be implemented.

Acknowledgments

This work was carried out as part of the Ph.D. study project "Monitoring Animal Wellbeing" at Technical University of Denmark. The project is partly funded by Danish Meat Research Institute. The authors thank to Pia Brandt for collecting and annotating the videos.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.compag.2015.09.021.

References

Ahrendt, P., Gregersen, T., Karstoft, H., 2011. Development of a real-time computer vision system for tracking loose-housed pigs. Comput. Electron. Agric. 76, 169–174.
Altunbasak, Y., Mersereau, R., Patti, A., 2003. A fast parametric motion estimation algorithm with illumination and lens distortion correction. IEEE Trans. Image Process. 12, 395–408.
Ardekani, R., Greenwood, A., Peichel, C., Tavare, S., 2013. Automated quantification of the schooling behaviour of sticklebacks. EURASIP J. Image Video Process. 2013, 61.
Bailey, D.G., 2011. Blob detection and labelling. In: Design for Embedded Image Processing on FPGAs. John Wiley & Sons (Asia) Pte Ltd., pp. 343–375 (Chapter Blob Detection and Labelling).
Boiman, O., Irani, M., 2007. Detecting irregularities in images and in video. Int. J. Comput. Vision 74, 17–31.
Brandt, P., Rousing, T., Herskin, M., Aaslyng, M., 2013. Identification of post-mortem indicators of welfare of finishing pigs on the day of slaughter. Livest. Sci. 157, 535–544.
Breslav, M., Fuller, N., Betke, M., 2012. Vision system for wing beat analysis of bats in the wild. In: International Conference on Pattern Recognition.
Broom, D., 2005. The effects of land transport on animal welfare. Rev. Sci. Tech.-Office Int. épizooties 24, 683.
Caja, G., Hernandez-Jover, M., Conill, C., Garin, D., Alabern, X., Farriol, B., Ghirardi, J., 2005. Use of ear tags and injectable transponders for the identification and traceability of pigs from birth to the end of the slaughter line. J. Anim. Sci. 83, 2215–2224.
Cangar, O., Leroy, T., Guarino, M., Vranken, E., 2008. Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis. Comput. Electron. Agric. 64, 53–60.
Costa, A., Ismayilova, G., Borgonovo, F., Viazzi, S., Berckmans, D., Guarino, M., 2014. Image-processing technique to measure pig activity in response to climatic variation in a pig barn. Anim. Prod. Sci. 54, 1075–1083.
Dawkins, M.S., Cain, R., Roberts, S.J., 2012. Optical flow, flock behaviour and chicken welfare. Anim. Behav. 84, 219–223.
Dawkins, M.S., Lee, H.-j., Waitt, C.D., Roberts, S.J., 2009. Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait. Appl. Anim. Behav. Sci. 119, 203–209.
Doane, D.P., 1976. Aesthetic frequency classifications. Am. Stat. 30, 181–183.
Gronskyte, R., Clemmensen, L., Hviid, M., Kulahci, M., 2015. Monitoring pigs' movement at the slaughterhouse using modified angular histograms. Biosyst. Eng., unpublished results.
Guo, Y., Zhu, W., Jiao, P., Chen, J., 2014. Foreground detection of group-housed pigs based on the combination of mixture of gaussians using prediction mechanism and threshold segmentation. Biosyst. Eng. 125, 98–104.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, vol. 2. Springer, New York.
Hendriks, C., Yu, Z., Lecocq, A., Bakker, T., Locke, B., Terenius, O., 2012. Identifying all individuals in a honeybee hive-progress towards mapping all social interactions. In: International Conference on Pattern Recognition.
Hu, W., Tan, T., Wang, L., Maybank, S., 2004. A survey on visual surveillance of object motion and behaviors. IEEE Trans. Syst., Man, Cybernet., Part C: Appl. Rev. 34, 334–352.
Ibrahim, N., Mokri, S., Siong, L., Mustafa, M., Hussain, A., 2010. Snatch theft detection using low level features. In: Proceedings of the World Congress on Engineering, vol. 2.
Ihaddadene, N., Djeraba, C., 2008. Real-time crowd motion analysis. In: 19th International Conference on Pattern Recognition, 2008. ICPR 2008, pp. 1–4.
Johnson, R., Freund, J., Miller, I., 2010. Miller and Freund's Probability and Statistics for Engineers. Prentice Hall.
Kashiha, M., Bahr, C., Ott, S., Moons, C., 2013. Automatic identification of marked pigs in a pen using image pattern recognition. Comput. Electron. Agric. 93, 111–120.
Kashiha, M., Bahr, C., Ott, S., Moons, C.P., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2014. Automatic weight estimation of individual pigs using image analysis. Comput. Electron. Agric. 107, 38–44.

Kehlbacher, A., Bennett, R., Balcombe, K., 2012. Measuring the consumer benefits of improving farm animal welfare to inform welfare labelling. Food Policy 37, 627–633.

Kongsro, J., 2013. Development of a computer vision system to monitor pig locomotion. Open J. Anim. Sci. 3, 254.

Kratz, L., Nishino, K., 2009a. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, pp. 1446–1453.

Kratz, L., Nishino, K., 2009b. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, pp. 1446–1453.

MATLAB, 2015. version 8.5.0 (R2015a). Natick, Massachusetts: The MathWorks Inc.

Mehran, R., Oyama, A., Shah, M., 2009. Abnormal crowd behavior detection using social force model. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, pp. 935–942.

Napolitano, F., Girolami, A., Braghieri, A., 2010. Consumer liking and willingness to pay for high welfare animal-based products. Trends Food Sci. Technol. 21, 537–543.

Ng, M.L., Leong, K.S., Hall, D., Cole, P.H., 2005. A small passive UHF RFID tag for livestock identification. In: IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, 2005. MAPE 2005, vol. 1, pp. 67–70.

Oczak, M., Viazzi, S., Ismayilova, G., Sonoda, L.T., Roulston, N., Fels, M., Bahr, C., Hartung, J., Guarino, M., Berckmans, D., 2014. Classification of aggressive behaviour in pigs by activity index and multilayer feed forward neural network. Biosyst. Eng. 119, 89–97.

Perko, R., Schnabel, T., Fritz, G., Almer, A., Paletta, L., 2013. Airborne based high performance crowd monitoring for security applications. In: Image Analysis, pp. 664–674.

Prola, L., Perona, G., Tursi, M., Mussa, P.P., 2010. Use of injectable transponders for the identification and traceability of pigs. Ital. J. Anim. Sci. 9, 35.

Roberts, S.J., Cain, R., Dawkins, M.S., 2012. Prediction of welfare outcomes for broiler chickens using bayesian regression on continuous optical flow data. J. Roy. Soc. Interface.

Soille, P., 2003. Morphological Image Analysis: Principles and Applications. Springer Science & Business Media.

Spink, A., Tegelenbosch, R., Buma, M., Noldus, L., 2001. The EthoVision video tracking system – a tool for behavioral phenotyping of transgenic mice. Physiol. Behav. 73, 731–744.

Tillett, R., Onyango, C., Marchant, J., 1997. Using model-based image processing to track animal movements. Comput. Electron. Agric. 17, 249–261.

Tu, G.J., Karstoft, H., Pedersen, L.J., Jørgensen, E., 2013. Foreground detection using loopy belief propagation. Biosyst. Eng. 116, 88–96.

Tuytelaars, T., Mikolajczyk, K., 2008. Local invariant feature detectors: a survey. Found. Trends Comput. Graph. Vision 3, 177–280.

Verbeke, W.A., Viaene, J., 2000. Ethical challenges for livestock production: meeting consumer concerns about meat safety and animalwelfare. J. Agric. Environ. Ethics 12, 141–151.

Warriss, P., Brown, S., Adams, S., Corlett, I., 1994. Relationships between subjective and objective assessments of stress at slaughter and meat quality in pigs. Meat Sci. 38, 329–340.

Appendix D

Appendix 4

R. Gronskyte, L. Clemmensen, M. Kulahci, "Online Monitoring of Crowds", Submitted to: Computer Vision and Image Understanding Journal.

Online Monitoring of Crowds

Ruta Gronskyte[a,*], Line Harder Clemmensen[a], Murat Kulahci[a,b]

[a] *DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark*
[b] *Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden*

Abstract

We present a procedure for monitoring crowd movements in videos based on continuous decision values (CDV) obtained from the classification of individual frames. The majority of classification methods performs the final classification based on continuous values (probabilities, scores, etc.), which we propose to use for the continuous monitoring by applying the principles of statistical process control. The classification method, used as an example, is a support vector machines (SVM), and the continuous SVM classification decision values are used for the monitoring. The SVM CDV are noisy and highly auto-correlated. Thus, they are smoothed using locally weighted scatterplot smoothing and events are detected using a cumulative sum (CUSUM) control chart. The total classification rate is increased from 95.8% to 99.8% by using continuous monitoring instead of discrete SVM. The classification rate of frames with an abnormal event is increased from 77.3% to 99.3%. The proposed procedure also allows for the detection of the events on average up to 139 frames earlier, than using the discrete SVM classification approach. The method is applied to monitor a pig herd while the pigs are being unloaded from the trucks. Events such as tripping and stepping on each other are of interest.

Keywords: Abnormal behavior; Continuous monitoring; Crowd monitoring; Event detection; Statistical process control; Support Vector Machines.

*Coresponting author:
Email address: rgro@dtu.dk (Ruta Gronskyte)
URL: www.dtu.dk (Ruta Gronskyte)

1. Introduction

Nowadays video surveillance is used in a large variety of areas to monitor humans, vehicles, and animals. Human [1, 2] and vehicle [3, 4] monitoring is used to ensure public safety, prevent undesired events and ensure the steady flow of traffic. Analysis of animal movement is a useful tool to ensure animal welfare [5, 6, 7, 8], understand the movement of animals [9] and monitor laboratory animal behavior [10]. The broad application area has attracted a great deal of research interest. As a result, a large number of methods have been proposed.

In some cases individual objects are identified, tracked and their motion is analyzed [11, 12, 13, 14, 15]. Individuals are rarely identified and tracked in large dense crowds which are common in parades [16], music events, religious gatherings [17], subways [2], etc. The methods to analyze these cases are classified as non-computer vision or computer vision based methods [18]. The former uses principles of psychology and sociology. The latter approach extracts features from the frame that are sometimes modeled and then used for classification. The analysis of a dense crowd using computer vision methods usually takes the following steps: (1) feature extraction, (2) motion and behavior modeling and (3) classification. The most important features are: crowd density, location, speed and movement direction. The density of a crowd can be estimated from the number of foreground pixels [19] or by applying pattern recognition techniques for a surveyed scene [20, 21] or combining both approaches. In most cases, the relative speed is estimated using optical flow (OF). Other measurements, such as kinetic energy, motion direction, distance, velocity, density, etc. [22, 23, 24, 25, 26], are also used. Multiple crowd models such as hidden Markov model [2, 27], a social force model [1], a social entropy measure [28] or a mixture of dynamic textures [29], have been proposed in the literature. Some methods model crowd movement using information from a single frame, while others use spatial-temporal informtion. The classification of a single frame is most commonly made using a thresholding approach [2, 29, 24, 25], using support vector machines (SVM) [28, 22, 23], K-means [1] or neural networks [26]. Although

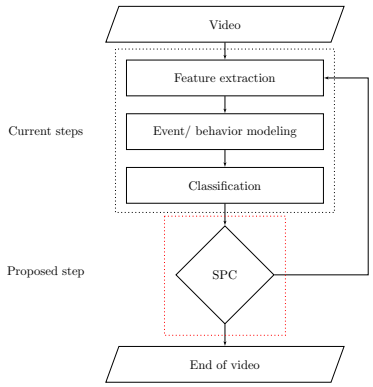


Figure 1: Flowchart of the proposed crowd monitoring approach.

some authors have built models and extracted features based on several frames, classification is usually done for a single frame disregarding the classification of the previous frames.

In most cases, human motion is continuous and transitions from one action or event to another. This implies that actions and events do not have a clear start or end. In this paper, we propose a procedure that allows for monitoring motion continuously. It is achieved by adding one more step, based on statistical process control (SPC), to the framework of dense crowd monitoring. A flowchart for crowd analysis is presented in Figure 1. Krausz et al. ([30]) employed an SPC control chart to monitor a histogram symmetry measure to detect crowd congestion. In this paper, we show how multiple features can be used to detect events by employing continuous classification values of SVM.

1.1. Statistical Process Control

Since late 1920's, the principles of SPC have been used to improve and ensure the quality of the processes. In the framework of SPC, any process exhibits two types of variability: *common cause* and *assignable cause*. The former variability is inherent process variability and exists in most stable systems. For example, in crowd movement monitoring, a source of such variability is from errors in the

OF estimation, feature extraction or classification. It is difficult to determine if one or more of these steps cause a classification error. For example, several OF methods are compared in [31] and it is concluded that most methods still have problems in the OF estimation of objects with complex structure and fast motion. Thus, errors can occur in the first step of motion analysis and affect subsequent steps. If a process only exhibits common cause variation, then this process is considered to be "in-control". In the analyzed case, the process is in-control when no abnormal behavior or undesired events occur in the video. The latter variability occurs when changes in the process appare due to assignable causes. Such causes can be any kind of unexpected event or abnormal behavior. Further in this paper we will refer to an "event" as any action of interest. A process is considered to be "out-of-control" when there is an assignable cause present in the system. In SPC, process monitoring is done with a control chart on which the quality characteristics of interest is plotted. A process is in-control when the quality characteristic remains between the predefined control limits and out-of-control when it exceeds these limits. The control limits define the acceptable level of common cause variation around the expected process mean. In this paper, we present three contributions, which combine the principles of SPC and crowd monitoring using video recordings:

- We propose using SPC as a fourth step in crowd monitoring. The outcome is an easily interpretable control chart, which decreases the misclassification rate and enhances the early detection of events.
- For the SPC charts, we propose to use continuous classification decision values (CDV) instead of categorical classes. Most classification methods assign a class based on a probability, score or any other type of decision values. The CDV capture the smooth transition between events, allowing events to be detected in the early stages of transition.
- Care must be taken when selecting historical in-control data for training the classifier and computing the in-control limits for the SPC chart. Manual annotation of highly auto-correlated video frames is labor intensive

and an error-prone task. In this paper, we propose some rules for frame selection and annotation.

1.2. Data

The proposed procedure is applied to monitor pig herd motion at the slaughterhouses when the pigs are unloaded from a truck. During the transportation, pigs are sedentary thus can be slow when moving out of the truck. Personnel uses specially designed sticks to encourage pigs to move faster. Some pigs become agitated and may cause a stampede. This in turn results in undesirable events such as tripping and stepping on each other. The aim of this case study is to monitor the unloading and prevent undesirable events, such as tripping and stepping on each other.

The videos are recorded using GoPro HERO2 (©2013 Woodman Labs, Inc) using 30 frames per second rate. The frame size is 1920 pixels in height and 1080 pixels in width. In a truck, pigs are divided into pens. A "clip" represents a video section where one pen is unloaded. In total 15 clips are used for the analysis. In each clip, one or two sections of data are annotated as in-control or out-of-control. Only those sections where behavior is clearly understood are annotated. As a result, 11 sections are declared as out-of-control, and 7 sections are declared as in-control. In total, 13,518 frames are used. In this paper, only the continuous movement monitoring based on SPC is discussed. The image analysis and event detection using discrete SVM classification in the presented case study are discussed in [7].

The following sections discuss the individual steps of continuous motions monitoring. The results of SVM classification and continuous monitoring based on SPC are compared, and presented in section Results.

2. Methodology

In this section, the individual steps of the crowd monitoring based on CDV are presented. As discussed in the previous section, several classification methods are used to detect events in crowds. Our method of choice is SVM and

in-depth motivation for the method can be found in [7]. We propose to use SVM Classification Decision Values instead of classification itself to monitor motion. CDV are noisy thus smoothing must be applied. We show that good results can be achieved using Locally weighted scatterplot smoothing (LOWESS). The smoothed values are monitored using a Cumulative Sum Control Chart (CUSUM).

2.1. Support Vector Machines

Support Vector Machines (SVM) [32, p. 417-455] is a supervised machine learning algorithm that classifies data into two classes. Support vector machines construct hyper-planes in higher order space such that the distance to the hyperplane and the nearest training data points from two classes is the largest. Let's define a training dataset consisting of a finite set of 2D points$(x_i; y_i)$ where $x_i \in \mathbb{R}$ and $y_i \in \{-1; 1\}$ represents two class labels and $i = 1, ..., N$. The separating line between two classes is defined as:

$$\{x : f(x) = x^T\beta + \beta_0 = 0\} \tag{1}$$

where β is a unit vector. The classification rule for the two classes is:

$$G(x) = sign[x^T\beta + \beta_0] \tag{2}$$

where we can find a function $f(x) = x^T\beta + \beta_0$ such that $y_i f(x_i) > 0 \ \forall i$. The SVM is defined as an optimization problem, where the data can be separated with the largest margin $\frac{1}{\|\beta\|}$ between the two groups. The optimization problem for non-overlapping classes can then be written as:

$$\begin{gathered} \arg\max_{\beta,\beta_0} \|\beta\| \\ \text{subject to } y_i(x_i^T\beta + \beta_0) \geq 1, i = 1, ..., N \end{gathered} \tag{3}$$

In most of the cases, the two classes cannot be perfectly sepereated. Let ξ_i be a measure called *slack* that defines the degree of misclassification of data point x_i:

$$y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i \tag{4}$$

Then the optimization problems is a trade-off between a large margin and a small misclassification:

$$\arg\min_{\beta,\beta_0}\|\beta\| \text{ subject to } \begin{cases} y_i(x_i^T\beta+\beta_0) \geq 1-\xi_i, \forall i \\ \xi_i \geq 0, \sum \xi_i \leq C \end{cases} \tag{5}$$

where C is a constant. In some cases, a better classification can be achieved by representing data in higher dimensions. Let $h(x)$ be a nonlinear basis function that transforms data into an extended feature space of higher dimensions, than $\hat{f}(x) = h(x)^T\beta + \beta_0$, and the classification rule is:

$$G(x) = sign[h(x)^T\beta + \beta_0] \tag{6}$$

In this paper, the values of $\hat{f}(x_i)$ will be referred to as SVM classification decision value (SVM CDV) or simply CDV. The three most popular kernel types used for data transformation are: (1) d^{th} polynomial, (2) radial basis and (3) neural network kernels. Cross-validation is used to find an optimal transformation function as well as a set of hyper-parameters. In this paper, the leave-one-clip-out cross-validation, which used 14 clips for training and one clip for testing, was used for indentification of the parameters. The advantages of using SVM are that it avoids over-fitting, performs well using a low number of samples and a high number of variables, and is not affected by cross-correlation.

2.2. Locally weighted scatterplot smoothing

Local regression (LOESS) or Locally Weighted Scatterplot Smoothing (LOWESS) ([32, p. 191-198], [33, p. 309 - 376]) is a non-parametric regression method used to smooth scatter plots. Linear and non-linear least square regression models are fitted locally for data span of size N. The fitted functions over this span can be: constant, linear or a low-order polynomial. The most common polynomials are of order $d = 1$ or 2. In addition, the weights can be chosen to include the data from a span of N observations with different proportions. The LOESS solves the optimization problem at each point x_0:

$$\max_{\alpha(x_0),\beta_j(x_0),j=1,\ldots,d} \sum_{i=1}^{N} K_\lambda(x_0,x_i)[y_i - \beta_0(x_0) - \sum_{j=1}^{d}\beta_j(x_0)x_i^j]^2 \tag{7}$$

where $K_\lambda(x_0, x_i)$ is a weight function, $(x_0; y_0)$ data point from the span and $\beta_0(x_0) + \sum_{j=1}^{d} \beta_j(x_0) x_i^j$ polynomial of order d.

The advantage of the LOESS is that the smoothing is not based on a single model. If a single model approach does not provide reasonable smoothing due to the process dynamics, then LOESS can be employed. However, LOESS model can not be presented analytically in most cases.

2.3. Cumulative Sum Chart

The CUSUM chart ([34, p. 408 - 419] is very effective in detecting small shifts in a process. One more advantage of this chart is that it incorporates a sequence of observations in decision making. For an in-control mean of μ_0 CUSUM statistics for variable z_i is given as:

$$\begin{aligned} C_i^+ &= max[0, z_i - (\mu_0 + K) + C_i^+] \\ C_i^- &= max[0, (\mu_0 + K) - z_i + C_i^-] \end{aligned} \tag{8}$$

where K is the slack variable introduced for reducing the number of false alarms and with the starting values $C_i^+ = C_i^- = 0$. The C_i^+ monitors the deviations that are above the process mean μ_0, while C_i^- deviations that are below. The process is declared as "out-of-control", when C_i^- or C_i^+ cross a pre-definde *threshold* H. The optimal K and H values are usually chosen by compromising between the false alarm rate and the early detection of an out-of-control situation.

2.4. Phases I and II

The statistical process control is carried out in two phases. [34, p. 198-199]. In *phase I*, which is off-line, historical data from an in-control process (training data) is collected, and the parameters of the control chart are estimated. The control limits are estimated to confirm that collected historical data are indeed in-control. The data points that are detected as out-of-control have to be inspected and removed if necessary. A new chart based on cleaned data is established and used in *phase II* for on-line monitoring.

The most commonly used structure for control charts is of the form:

$$\begin{aligned} UCL &= \mu_w + L\sigma_w \\ \text{Central line} &= \mu_w \\ LCL &= \mu_w - L\sigma_w \end{aligned} \tag{9}$$

where UCL and LCL are the upper and lower control limits respectively, μ_w and σ_w are process mean and standard deviation respectively, and L is a user defined parameter to define the band within which the process is deemed in-control. Sufficient amount of data in *phase I* is needed for a proper estimation of the parameters of the control chart. Some guidelines for univariate and multivariate data are provided in [34].

2.5. Types of errors

There are two types of errors [34, p. 182-193] related to declaring the process "in-control" or "out-of-control":

- *Type I error* indicates the probability of declaring a process in a state of out-of-control when it actually is in-control.
- *Type II error* indicates the probability of declaring the process in a state of in-control when it actually is out-of-control.

It is common to use the magnitudes of these errors to compare the performances of the charts as well as to choose the control limits for a chart. This is usually done by using the *average run length ARL*. The ARL_0 for an in-control process is defined as:

$$ARL_0 = \frac{1}{p} \tag{10}$$

where p is the *type I error*. If the *type I error* is fixed in *phase I*, then ARL_0 of two charts can be compared during *phase II*. The average run-length for an out-of-control process ARL_1 is defined as:

$$ARL_1 = \frac{1}{1-\beta} \tag{11}$$

where β is the probability of not detecting an out-of-control situation on the first subsequent sample, i.e. Type II error. It is expected that an effective control chart has the in-control ARL_0 as large as possible while out-of-control ARL_1 as small as possible.

3. The proposed procedure

For the majority of the proposed methods for crowd monitoring and event detection, the approach is based on classification. In this paper, we propose to add one more step to the well-established crowd monitoring framework (Figure 1). This additional step decreases the number of misclassifications, increases the early detection count for events and provides a real-time overview of whether actions taken to control a crowd is effective. We use the term "early detection count" (EDC) to indicate the number of observations declared as out-of-control before an actual event is observed. Also, the proposed method results in a decision that includes information of multiple consecutive frames hence taking into account the history.

In this case study, our classification method of choice is SVM, but the proposed procedure can accommodate any classification method that can provide a continuous decision value. The SVM CDVs are highly auto-correlated due to high sampling frequency and noisy, thus smoothing must first be applied. A sample clip of CDVs is presented in Figure 2. We propose to use LOESS with a fixed span N to smooth out the noise. The last step in the procedure is the establishment of a CUSUM chart.

The SVM continuous decision value is expected to be positive for in-control processes and negative for out-of-control processes. Therefore only C_i^- in equation 8 is of interest in our case. To establish continuous online monitoring, three parameters have to be chosen: (1) the span for LOESS smoothing, (2) $(\mu_0 + K)$ for the CUSUM chart in Equation 8 and (3) the threshold H. The parameters are identified using two approaches: (1) by only optimizing the total classification rate or EDC and (2) based on the SPC procedure for choosing

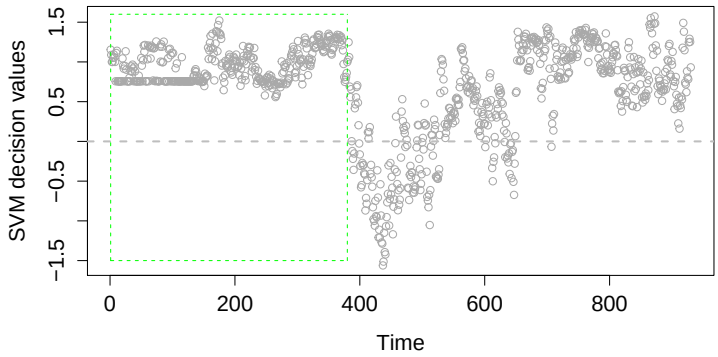


Figure 2: A sample of an in-control clip. The green box represents the data section that is used for training as the in-control class. The horizontal gray line indicates the standard SVM decision threshold. All observations above the line are in-control frames and all below are out-of-control.

control chart parameters by fixing the *Type I errors*. For that we consider a systematic approach of varying the values of the these parameters within the following ranges: $N = 5, 10, 20, \ldots, 60$, $(\mu_0 + K) = 0, 0.05, 0.1, \ldots, 0.6$ and $H = 0.001, 0.004, \ldots, 0.01, 0.03 \ldots, 0.09, 0.1$. It should be noted that a span of $N = 30$ represents approximately 1 second of a video.

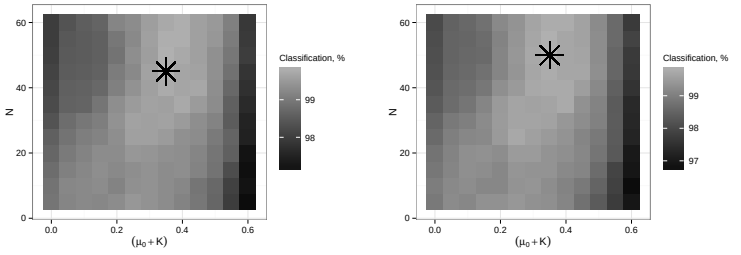
As we discussed in the Introduction, a motion is continuous and does not start or end instantaneously. One must be very careful when selecting historical in-control data from videos. [35] compares spatial, temporal and behavioral labeling of human action and concludes that errors accrue at each level of the annotation. For example, the majority of temporal disagreements are at the beginning and the end of an event. The various errors appear due to the annotator's understanding and interpretation of the scene. In our experience, the complexity of annotation increases with the increased number of surveyed objects as well as with the increased number of action labels. Thus, we propose

using only those frames where motion is clearly present and confirmed.

4. Results

4.1. Parameter identification by only optimizing total classification or early detection count

In this subsection, the optimal parameters are identified to optimize: (1) the total classification rate and (2) the EDC. Two sets of parameters are determined that achieve the highest total classification rate *set 1*: $N = 45$, $(\mu_0 + K) = 0.35$ $H = 0.05$ and *set 2*: $N = 50$, $(\mu_0 + K) = 0.35$ $H = 0.001$. The identification of parameters is shown in Figure 3.



(a) $H = 0.05$. (b) $H = 0.001$.

Figure 3: Parameter identification aiming to optimize the total classification rate. The highest classification rate is achieved with two sets of parameters.

The set of optimal parameters that optimize the EDC is *set 3*: $N = 35$, $(\mu_0 + K) = 0.6$ $H = 0.001$. The identification of the parameters is presented in Figure 4.

The performance of three identified parameter sets is compared. The total, in-control and out-of-control classification rates are compared as well as EDC. The Shapiro-Wilk normality test indicates that the difference between SVM and CUSUM chart classification rates are not Gaussian distributed. Thus, Wilcoxon singed-rank test is used to test the significance of the improved performance. A comparison of SVM classification and CUSUM chart is presented in Table 1.

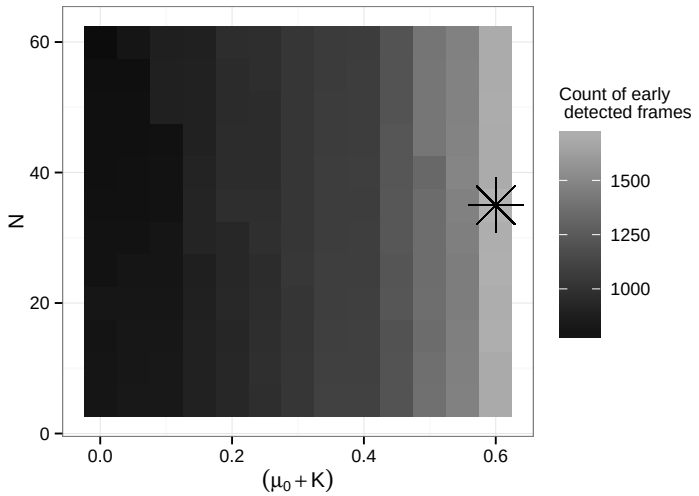


Figure 4: Parameter identification aiming to optimize the EDC. The earliest event detection is achieved with parameters: $N = 35$, $(\mu_0 + K) = 0.6$ and $H = 0.001$.

The comparison indicates that there is no significant improvement of total and in-control frame classification. SVM classifies correctly most in-control frames. Thus, significant improvement is not possible. However, the CUSUM chart significantly improves out-of-control frame classification with all sets of parameters.

The comparison of EDC of all three sets is presented in the Figure 5. The *set 3* parameters detects events earlier than *sets 1* & *2*.

The Wilcoxon signed-rank test is performed (Table 2) to compare the EDCs. The comparison indicates that the CUSUM chart detects events earlier than SVM classification with all sets of parameters. The CUSUM chart, with *set 3* parameters, can detect on average 139 frames earlier than SVM.

The results suggest that to obtain a better classification rate, a larger smoothing span N for LOESS must be chosen. However, a large span will add lag to the changes in the dynamic process, thus decreasing the EDC. The $(\mu_0 + K)$

		Set 1		Set 2		Set 3	
	SVM	CUSUM	p-val	CUSUM	p-val	CUSUM	p-val
Total	95.6 %	99.8 %	0.998	99.8%	0.998	97.7%	0.966
In	99.8 %	100 %	0.977	100%	0.977	96.4 %	0.186
Out	74.3%	99.3%	0.002	99.4%	0.002	99.9%	0.002

Table 1: Comparison of SVM and CUSUM classification rates using three CUSUM parameter sets. Only for the out-of-control case is the CUSUM classification rate significantly ($p < 0.05$) better than the SVM.

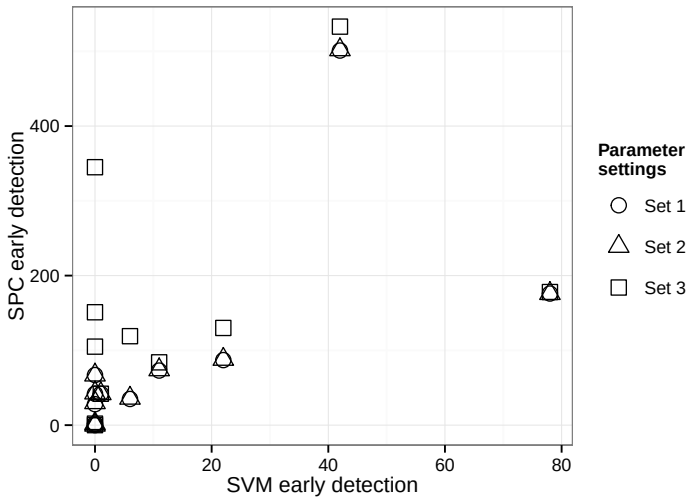


Figure 5: EDCs comparison of three parameter sets.

value closer to the in-control process mean must be selected to obtain better EDC. A sample clip using different parameters is presented in Figure 6.

4.2. Parameter identification using SPC approach

The first step in parameter identification using SPC approach is to fix the false alarm rate α. In this case, we fix the false alarm rate to the SVM mis-

Set	p-value	Number of frames detected earlier than SVM
1	0.006	81.18
2	0.006	81.81
3	0.006	139

Table 2: EDCs comparison of three CUSUM parameter sets. For all parameter sets, events are detected earlier compared to the standard approach ($p < 0.05$).

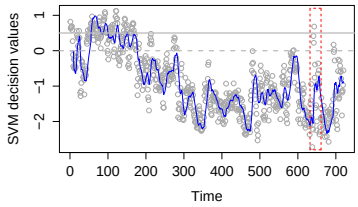
classification rate of annotated in-control frames, which is $\alpha = 0.1705\%$. For different combinations of N and $(\mu_0 + K)$ a minimal value of H is identified such that the in-control classification rate of annotated frames using a CUSUM chart is 99.8295%. Thirty-five sets of parameters N, $(\mu_0 + K)$ and H are identified. The pairs of parameters are shown in Figure 7. The EDC and out-of-control frame classification rates of SVM and CUSUM chart are compared. The results for EDC are also presented in Figure 7.

The CUSUM chart significantly improves the EDC in comparison to SVM at a significance level of 0.05, when the false alarm rate is fixed to $\alpha = 0.1705\%$. The CUSUM chart can detect events from 59 to 82 frames earlier, which is approximately $2 - 2.7$ seconds.

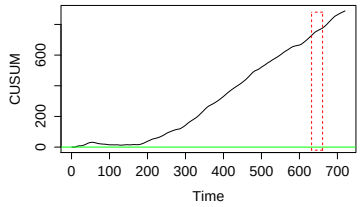
The combinations of identified parameters (N, $(\mu_0 + K)$ and H) achieve between $98\% - 100\%$ correct out-of-control frame classification rate, using the CUSUM chart. The Wilcoxon test indicates that CUSUM chart improves the out-of-control classification by around 25% at a significance level 0.05 as presented in Figure 7c.

5. Discussion

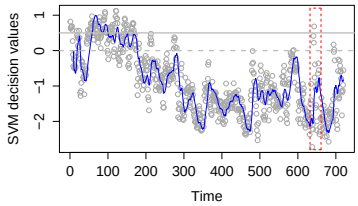
The above-presented results indicate that the classification rate and EDC can be improved using a CUSUM chart for smoothed CDVs. Three parameters have to be chosen in the proposed procedure: (1) the smoothing span N, (2) the $(\mu_0 + K)$ and (3) the decision interval H.



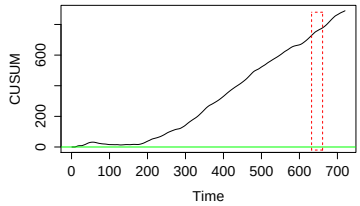
(a) CDV smoothed using *set 1* parameters.



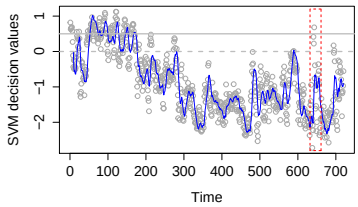
(b) CUSUM chart of 6a.



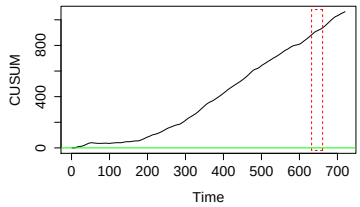
(c) CDV smoothed using *set 2* parameters.



(d) CUSUM chart of 6c.

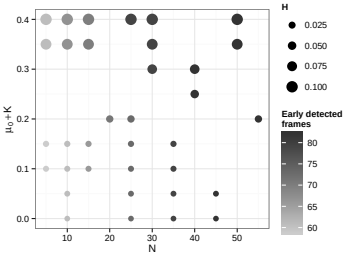


(e) CDV smoothed using *set 3* parameters.

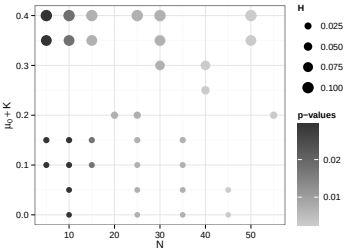


(f) CUSUM chart of 6e.

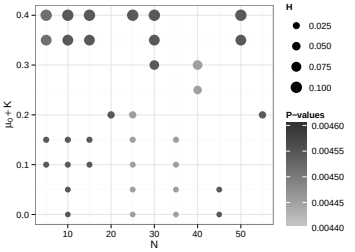
Figure 6: A sample clip smoothed using different span $N = 35, 45, 50$ is presented on left. The CUSUM chart of the smoothed clips is presented on the right. The red boxes represent annotated events, the gray dash line represents SVM decision threshold, the solid gray line represents $(\mu_0 + K)$, gray points represents CDV and blue points smoothed CDV.



(a) Mean of early predicted frames.



(b) Wilcoxon test for early predicted frames.



(c) Wilcoxon test for out-of-control frame classification.

Figure 7: Comparison of SVM and CUSUM chart. The p values indicate that a CUSUM chart detects events significantly earlier and classifies significantly more frames with events correctly with a 0.05 significance level. The CUSUM chart can detect events from 59 to 82 frames earlier than SVM.

The analysis of different parameter combinations based on our dataset reveals that the most important parameter is the smoothing span N. A large span generates heavier smoothing, thus removing more noise and increasing the classification rate. However, choosing a span too high does not improve the EDC.

The value of $(\mu_0 + K)$ highly depends on the chosen N value. If a large N is chosen, then the process is smooth and a value close to the in-control process mean can be chosen for $(\mu_0 + K)$. Choosing a small span and large $(\mu_0 + K)$

will lead to an increased false alarm rate. If the values of N and $(\mu_0 + K)$ are chosen appropriately, the decision interval H can be very small, leading to early detection of out-of-control situations.

6. Conclusion

Nowadays, video surveillance is used in many fields. In this paper, the main interest is crowd monitoring. The classic approach for detecting any kind of event is to use classification methods for crowd models or extracted features from single or multiple frames. In this way, the motion is discretized, and it has a clear beginning and an end. By treating motion in this way, important information is lost. In this paper, we propose applying an additional step to the crowd monitoring framework and analyze motion based on continuous decision values.

The majority of classification methods make a decision based on scores, probabilities or decision values that are continuous. By monitoring these continuous values, more information about the process is attained. In this paper, we analyze the support vector machines' (SVM) classification decision values (CDVs) of pig herd movement. The values are highly auto-correlated and noisy, thus smoothing is applied. Good smoothing results are obtained using local polynomial smoothing (LOESS), and the process is monitored using a cumulative sum (CUSUM) chart.

The results indicate that the procedure can outperform SVM classification alone. The total SVM classification rate of 95.8% can be improved up to 99.8%. The early event detection count can be on average improved by 139 frames. The proposed methodology can be used for any classifier that is based on a continuous decision variable. In that sense, there potentially exist many other possibilities of applications of the proposed methodology in continuous process surveillance through image analysis.

Acknowledgment

This work was carried out as part of the Ph.D. study project Monitoring Animal Wellbeing at the Technical University of Denmark. The project is partly funded by the Danish Meat Research Institute. The authors thank to Pia Brandt and Marchen Sonja Hviid for collecting and annotating the videos.

References

[1] R. Mehran, A. Oyama, M. Shah, Abnormal crowd behavior detection using social force model, in: Proceedings of 2009 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR), 2009, pp. 935–942. doi:10.1109/CVPR.2009.5206641.

[2] L. Kratz, K. Nishino, Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models, in: Proceedings of 2009 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR), 2009, pp. 1446–1453. doi:10.1109/CVPR.2009.5206771.

[3] J. Zhu, Y. Lao, Y. F. Zheng, Object tracking in structured environments for video surveillance applications, IEEE Transactions on: Circuits and Systems for Video Technology 20 (2) (2010) 223–235. doi:10.1109/TCSVT.2009.2031395.

[4] B. T. Morris, M. M. Trivedi, Learning, modeling, and classification of vehicle track patterns from live video, IEEE Transactions on: Intelligent Transportation Systems 9 (3) (2008) 425–437. doi:10.1109/TITS.2008.922970.

[5] M. S. Dawkins, H.-j. Lee, C. D. Waitt, S. J. Roberts, Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait, Applied Animal Behaviour Science 119 (3) (2009) 203–209. doi:10.1016/j.applanim.2009.04.009.

[6] R. Gronskyte, L. H. Clemmensen, M. S. Hviid, M. Kulahci, Monitoring pig movement at the slaughterhouse using optical flow and modified angular

histograms, Biosystems Engineering 141 (2016) 19–30. doi:10.1016/j.biosystemseng.2015.10.002.

[7] R. Gronskyte, L. H. Clemmensen, M. S. Hviid, M. Kulahci, Pig herd monitoring and undesirable tripping and stepping prevention, Computers and Electronics in Agriculture 119 (2015) 51–60. doi:10.1016/j.compag.2015.09.021.

[8] P. Brandt, T. Rousing, M. Herskin, M. Aaslyng, Identification of *post-mortem* indicators of welfare of finishing pigs on the day of slaughter, Livestock Science 157 (2) (2013) 535–544. doi:10.1016/j.livsci.2013.08.020.

[9] M. Breslav, N. W. Fuller, M. Betke, Vision system for wingbeat analysis of bats in the wild, in: Proceedings of 21st International Conference on: Pattern Recognition (PR), 2012.
URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.303.142&rep=rep1&type=pdf

[10] E. A. van Dam, J. E. van der Harst, C. J. ter Braak, R. A. Tegelenbosch, B. M. Spruijt, L. P. Noldus, An automated system for the recognition of various specific rat behaviours, Journal of Neuroscience Methods 218 (2) (2013) 214–224. doi:10.1016/j.jneumeth.2013.05.012.

[11] W. Hu, T. Tan, L. Wang, S. Maybank, A survey on visual surveillance of object motion and behaviors, IEEE Transactions on: Systems, Man, and Cybernetics, Part C: Applications and Reviews 34 (3) (2004) 334–352. doi:10.1109/TSMCC.2004.829274.

[12] A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, S. Pankanti, Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking, Signal Processing Magazine 22 (2) (2005) 38–51. doi:10.1109/MSP.2005.1406476.

[13] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, K.-H. Choi, A review on video-based human activity recognition, Computers 2 (2) (2013) 88–131. doi:10.3390/computers2020088.

[14] O. P. Popoola, K. Wang, Video-based abnormal human behavior recognitiona review, IEEE Transactions on: Systems, Man, and Cybernetics, Part C: Applications and Reviews 42 (6) (2012) 865–878. doi:10.1109/TSMCC.2011.2178594.

[15] M. Cristani, R. Raghavendra, A. Del Bue, V. Murino, Human behavior analysis in video surveillance: A social signal processing perspective, Neurocomputing 100 (2013) 86–97. doi:10.1016/j.neucom.2011.12.038.

[16] B. Krausz, C. Bauckhage, Loveparade 2010: Automatic video analysis of a crowd disaster, Computer Vision and Image Understanding 116 (3) (2012) 307–319. doi:10.1016/j.cviu.2011.08.006.

[17] D. Helbing, A. Johansson, H. Z. Al-Abideen, Dynamics of crowd disasters: An empirical study, Physical review E 75 (4) (2007) 046109. doi:http://dx.doi.org/10.1103/PhysRevE.75.046109.

[18] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, L.-Q. Xu, Crowd analysis: a survey, Machine Vision and Applications 19 (5-6) (2008) 345–357. doi:10.1007/s00138-008-0132-4.

[19] S. Velastin, J. Yin, A. Davies, M. Vicencio-Silva, R. Allsop, A. Penn, Automated measurement of crowd density and motion using image processing, in: Proceedings of 7th International Conference on: Road Traffic Monitoring and Control, 1994, pp. 127–132.
URL http://trid.trb.org/view.aspx?id=413500

[20] H. Rahmalan, M. S. Nixon, J. N. Carter, On crowd density estimation for surveillance, in: Proceedings of 2006 IET Conference on: Crime and Security, 2006, pp. 540–545(5). doi:10.1049/ic:20060360.

[21] S.-F. Lin, J.-Y. Chen, H.-X. Chao, Estimation of number of people in crowded scenes using perspective transformation, IEEE Transactions on: Systems, Man and Cybernetics, Part A: Systems and Humans 31 (6) (2001) 645–654. doi:10.1109/3468.983420.

[22] H. Liao, J. Xiang, W. Sun, Q. Feng, J. Dai, An abnormal event recognition in crowd scene, in: Proceedings of 6th International Conference on: Image and Graphics (ICIG), 2011, pp. 731–736. doi:10.1109/ICIG.2011.66.

[23] M. Haque, M. Murshed, Panic-driven event detection from surveillance video stream without track and motion features, in: Proceedings of 2010 IEEE International Conference on: Multimedia and Expo (ICME), 2010, pp. 173–178. doi:10.1109/ICME.2010.5583057.

[24] N. Ihaddadene, C. Djeraba, Real-time crowd motion analysis, in: Proceedings of 19th International Conference on: Pattern Recognition (ICPR), 2008, pp. 1–4. doi:10.1109/ICPR.2008.4761041.

[25] K. Goya, X. Zhang, K. Kitayama, I. Nagayama, A method for automatic detection of crimes for public security by using motion analysis, in: Proceedings of 5th International Conference on: Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2009, pp. 736–741. doi:10.1109/IIH-MSP.2009.264.

[26] N. Rasheed, S. A. Khan, A. Khalid, Tracking and abnormal behavior detection in video surveillance using optical flow and neural networks, in: Proceedings of 28th International Conference on: Advanced Information Networking and Applications (WAINA), 2014, pp. 61–66. doi:10.1109/WAINA.2014.18.

[27] E. L. Andrade, S. Blunsden, R. B. Fisher, Hidden markov models for optical flow analysis in crowds, in: Proceedings of 18th International Conference on: Pattern Recognition (ICPR), Vol. 1, 2006, pp. 460–463. doi:10.1109/ICPR.2006.621.

[28] S. Pathan, A. Al-Hamadi, B. Michaelis, Incorporating social entropy for crowd behavior detection using SVM, in: Advances in Visual Computing, Vol. 6453, 2010, pp. 153–162. doi:10.1007/978-3-642-17289-2_15.

[29] V. Mahadevan, W. Li, V. Bhalodia, N. Vasconcelos, Anomaly detection in crowded scenes, in: Proceedings of 2010 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR), 2010, pp. 1975–1981. doi:10.1109/CVPR.2010.5539872.

[30] B. Krausz, C. Bauckhage, Automatic detection of dangerous motion behavior in human crowds, in: Proceedings of 8th IEEE International Conference on: Advanced Video and Signal-Based Surveillance (AVSS), 2011, pp. 224–229. doi:10.1109/AVSS.2011.6027326.

[31] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. J. Black, R. Szeliski, A database and evaluation methodology for optical flow, International Journal of Computer Vision 92 (1) (2011) 1–31. doi:10.1007/s11263-010-0390-2.

[32] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, R. Tibshirani, The elements of statistical learning, Vol. 2, 2009.

[33] J. M. Chambers, T. J. Hastie, Statistical models in S, 1992.

[34] D. C. Montgomery, Introduction to statistical quality control, 2007.

[35] T. List, J. Bins, J. Vazquez, R. B. Fisher, Performance evaluating the evaluator, in: Proceedings of 2nd Joint IEEE International Workshop on: Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005, pp. 129–136. doi:10.1109/VSPETS.2005.1570907.

Bibliography

[1] H. J. Blokhuis, L. J. Keeling, A. Gavinelli, and J. Serratosa, "Animal welfare's impact on the food chain," *Trends in Food Science & Technology*, vol. 19, pp. S79–S87, 2008.

[2] B. Boogaard, S. Oosting, and B. Bock, "Elements of societal perception of farm animal welfare: A quantitative study in the netherlands," *Livestock Science*, vol. 104, no. 1, pp. 13–22, 2006.

[3] D. M. Broom, "Indicators of poor welfare," *British Veterinary Journal*, vol. 142, no. 6, pp. 524–526, 1986.

[4] E. Terlouw, C. Arnould, B. Auperin, C. Berri, E. Le Bihan-Duval, V. Deiss, F. Lefevre, B. Lensink, and L. Mounier, "Pre-slaughter conditions, animal stress and welfare: current status and possible future research," 2008.

[5] P. Brandt, T. Rousing, M. S. Herskin, and M. D. Aaslyng, "Identification of post-mortem indicators of welfare of finishing pigs on the day of slaughter," *Livestock Science*, vol. 157, no. 2, pp. 535–544, 2013.

[6] J. Choe and B. Kim, "Association of blood glucose, blood lactate, serum cortisol levels, muscle metabolites, muscle fiber type composition, and pork quality traits," *Meat Science*, vol. 97, no. 2, pp. 137–142, 2014.

[7] D. Broom, "The effects of land transport on animal welfare," *Revue scientifique et technique-Office international des épizooties*, vol. 24, no. 2, p. 683, 2005.

[8] A. Frost, C. Schofield, S. Beaulah, T. Mottram, J. Lines, and C. Wathes, "A review of livestock monitoring and the need for integrated systems," *Computers and Electronics in Agriculture*, vol. 17, no. 2, pp. 139–159, 1997.

[9] D. Fleet and Y. Weiss, *Handbook of Mathematical Models in Computer Vision*. Springer, 2006.

[10] D. A. Wedel and P. D. Cremers, *Stereo Scene Flow for 3D Motion Analysis*. Springer, 2011.

[11] J. J. Gibson, *The perception of the visual world*. 1950.

[12] B. D. Lucas, T. Kanade, and others, "An iterative image registration technique with an application to stereo vision.," in *Proceedings of International Conference on: Artificial Intelligence*, vol. 81, pp. 674–679, 1981.

[13] B. K. Horn and B. G. Schunck, "Determining optical flow," in *Proceedings of 1981 Technical Symposium East on: International Society for Optics and Photonics*, pp. 319–331, 1981.

[14] P. Soille, *Morphological image analysis: principles and applications*. 2003.

[15] Matlab, "Image processing toolbox (R2012b)," 2012.

[16] S. Brutzer, B. Hoferlin, and G. Heidemann, "Evaluation of background subtraction techniques for video surveillance," in *Proceedings of 2011 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR)*, pp. 1937–1944, 2011.

[17] M. Piccardi, "Background subtraction techniques: a review," in *Proceedings of 2004 IEEE International Conference on: Systems, man and cybernetics*, vol. 4, pp. 3099–3104, 2004.

[18] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," in *Proceedings of 2001 IEEE Conference on: Intelligent Transportation Systems*, pp. 334–339, 2001.

[19] Y. Altunbasak, R. M. Mersereau, and A. J. Patti, "A fast parametric motion estimation algorithm with illumination and lens distortion correction," *IEEE Transactions on: Image Processing*, vol. 12, no. 4, pp. 395–408, 2003.

[20] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, "A review on video-based human activity recognition," *Computers*, vol. 2, no. 2, pp. 88–131, 2013.

[21] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on: Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 34, no. 3, pp. 334–352, 2004.

[22] T. Ko, "A survey on behavior analysis in video surveillance for homeland security applications," in *Proceedings of 37th Applied Imagery Pattern Recognition Workshop*, pp. 1–8, 2008.

[23] A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, and S. Pankanti, "Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking," *Signal Processing Magazine*, vol. 22, no. 2, pp. 38–51, 2005.

[24] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, and L.-Q. Xu, "Crowd analysis: a survey," *Machine Vision and Applications*, vol. 19, no. 5-6, pp. 345–357, 2008.

[25] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," in *Proceedings of 2009 Conference on: Computer Vision and Pattern Recognition (CVPR)*, pp. 1446–1453, 2009.

[26] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in *Proceedings of 2010 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR)*, pp. 1975–1981, 2010.

[27] N. Ihaddadene and C. Djeraba, "Real-time crowd motion analysis," in *Proceedings of 19th International Conference on: Pattern Recognition.*, pp. 1–4, 2008.

[28] K. Goya, X. Zhang, K. Kitayama, and I. Nagayama, "A method for automatic detection of crimes for public security by using motion analysis," in *Proceedings of th International Conference on: Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'09)*, pp. 736–741, 2009.

[29] H. Liao, J. Xiang, W. Sun, Q. Feng, and J. Dai, "An abnormal event recognition in crowd scene," in *Proceedings of 6th International Conference on: Image and Graphics (ICIG)*, pp. 731–736, 2011.

[30] M. Haque and M. Murshed, "Panic-driven event detection from surveillance video stream without track and motion features," in *Proceeding of 2010 IEEE International Conference on: Multimedia and Expo (ICME)*, pp. 173–178, 2010.

[31] N. Rasheed, S. A. Khan, and A. Khalid, "Tracking and abnormal behavior detection in video surveillance using optical flow and neural networks," in *Proceedings of: Advanced Information Networking and Applications Workshops*, pp. 61–66, 2014.

[32] E. Andrade, S. Blunsden, and R. Fisher, "Hidden markov models for optical flow analysis in crowds," in *Proceedings of 18th International Conference on: Pattern Recognition (PR)*, vol. 1, pp. 460–463, 2006.

[33] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Computer Vision and Pattern Recognition*, pp. 935–942, 2009.

[34] S. Pathan, A. Al-Hamadi, and B. Michaelis, "Incorporating social entropy for crowd behavior detection using svm," in *Advances in Visual Computing*, vol. 6453, pp. 153–162, Springer Berlin Heidelberg, 2010.

[35] E. A. van Dam, J. E. van der Harst, C. J. ter Braak, R. A. Tegelenbosch, B. M. Spruijt, and L. P. Noldus, "An automated system for the recognition of various specific rat behaviours," *Journal of Neuroscience Methods*, vol. 218, no. 2, pp. 214–224, 2013.

[36] J. Rousseau, P. Van Lochem, W. Gispen, and B. Spruijt, "Classification of rat behavior with an image-processing method and a neural network," *Behavior Research Methods, Instruments, & Computers*, vol. 32, no. 1, pp. 63–71, 2000.

[37] T. Palmér and P. Petersson, "Automated tracking of motor behavior as a means to assess severity of symptoms in the 6-OHDA marmoset model of parkinsons disease," in *Proceedings of International Conference on: Pattern Recognition (PR)*, vol. 11, 2012.

[38] R. Ardekani, A. K. Greenwood, C. L. Peichel, and S. Tavaré, "Automated quantification of the schooling behaviour of sticklebacks," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, pp. 1–8, 2013.

[39] H. Dankert, L. Wang, E. D. Hoopfer, D. J. Anderson, and P. Perona, "Automated monitoring and analysis of social behavior in drosophila," *Nature Methods*, vol. 6, no. 4, pp. 297–303, 2009.

[40] C. L. L. Hendriks, Z. Yu, A. Lecocq, T. Bakker, B. Locke, and O. Terenius, "Identifying all individuals in a honeybee hive–progress towards mapping all social interactions," in *Proceedings of International Conference on: Pattern Recognition (PR)*, vol. 11, 2012.

[41] M. Breslav, N. W. Fuller, and M. Betke, "Vision system for wingbeat analysis of bats in the wild," in *Proceedings of International Conference on: Pattern Recognition (PR)*, vol. 11, 2012.

[42] Ö. Cangar, T. Leroy, M. Guarino, E. Vranken, R. Fallon, J. Lenehan, J. Mee, and D. Berckmans, "Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis," *Computers and Electronics in Agriculture*, vol. 64, no. 1, pp. 53–60, 2008.

[43] M. S. Dawkins, R. Cain, and S. J. Roberts, "Optical flow, flock behaviour and chicken welfare," *Animal Behaviour*, no. 1, pp. 219–223, 2012.

[44] M. S. Dawkins, H.-j. Lee, C. D. Waitt, and S. J. Roberts, "Optical flow patterns in broiler chicken flocks as automated measures of behaviour and gait," *Applied Animal Behaviour Science*, vol. 119, no. 3, pp. 203–209, 2009.

[45] M. Kashiha, C. Bahr, S. Ott, C. P. Moons, T. A. Niewold, F. O. Ödberg, and D. Berckmans, "Automatic identification of marked pigs in a pen using image pattern recognition," *Computers and Electronics in Agriculture*, vol. 93, pp. 111–120, 2013.

[46] P. Ahrendt, T. Gregersen, and H. Karstoft, "Development of a real-time computer vision system for tracking loose-housed pigs," *Computers and Electronics in Agriculture*, vol. 76, no. 2, pp. 169–174, 2011.

[47] R. Tillett, C. Onyango, and J. Marchant, "Using model-based image processing to track animal movements," *Computers and Electronics in Agriculture*, vol. 17, no. 2, pp. 249–261, 1997.

[48] J. E. Jackson, *A user's guide to principal components*. John Wiley & Sons, 1991.

[49] J. Yang, D. Zhang, A. F. Frangi, and J.-y. Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Transactions on: Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[50] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proceedings of 2005 IEEE computer society conference on: Computer vision and pattern recognition.*, vol. 1, pp. 947–954, 2005.

[51] E. L. Andrade, S. Blunsden, and R. B. Fisher, "Modelling crowd scenes for event detection," in *Proceedings of 18th International Conference on: Pattern Recognition ICPR*, vol. 1, pp. 175–178, IEEE, 2006.

[52] F. I. Bashir, A. A. Khokhar, and D. Schonfeld, "Segmented trajectory based indexing and retrieval of video data," in *Proceedings of 2003 International Conference on: Image Processing*, vol. 2, pp. II–623, 2003.

[53] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE Journal*, vol. 40, no. 8, pp. 1361–1375, 1994.

[54] J. MacGregor and T. Kourti, "Statistical process control of multivariate processes," *Control Engineering Practice*, vol. 3, no. 3, pp. 403–414, 1995.

[55] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*, vol. 2. Springer, 2009.

[56] H. Meng, N. Pears, and C. Bailey, "A human action recognition system for embedded computer vision application," in *Proceedings of 2007 IEEE Conference on: Computer Vision and Pattern Recognition (CVPR)*, pp. 1–6, 2007.

[57] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Proceedings of 8th IEEE International Conference on: Computer Vision (ICCV)*, vol. 2, pp. 688–694, IEEE, 2001.

[58] Y. Andreu, P. García-Sevilla, and R. A. Mollineda, "Face gender classification: A statistical study when neutral and distorted faces are combined for training and testing purposes," *Image and Vision Computing*, vol. 32, no. 1, pp. 27–36, 2014.

[59] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, no. 1-3, pp. 389–422, 2002.

[60] X. Wu, G. Liang, K. K. Lee, and Y. Xu, "Crowd density estimation using texture analysis and learning," in *Proceedings of 2006 IEEE International Conference on: Robotics and Biomimetics (ROBIO)*, pp. 214–219, IEEE, 2006.

[61] J. M. Chambers and T. J. Hastie, *Statistical models in S*. 1992.

[62] D. C. Montgomery, *Introduction to statistical quality control*. 6th ed., 2009.

[63] J. M. Lucas, "Counted data cusum's," *Technometrics*, vol. 27, no. 2, pp. 129–144, 1985.

[64] P. D. Bourke, "Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection," *Journal of Quality Technology*, vol. 23, no. 3, pp. 225–238, 1991.

[65] C. Lai, K. Govindaraju, and M. Xie, "Effects of correlation on fraction non-conforming statistical process control procedures," *Journal of Applied Statistics*, vol. 25, no. 4, pp. 535–543, 1998.

[66] U. Narayan Bhat and R. Lal, "Attribute control charts for markov dependent production processes," *IIE transactions*, vol. 22, no. 2, pp. 181–188, 1990.

[67] D. K. Shepherd, C. W. Champ, S. E. Rigdon, and H. T. Fuller, "Attribute charts for monitoring a dependent process," *Quality and Reliability Engineering International*, vol. 23, no. 3, pp. 341–365, 2007.

[68] C. A. Lowry and D. C. Montgomery, "A review of multivariate control charts," *IIE transactions*, vol. 27, no. 6, pp. 800–810, 1995.

[69] B. R. Bakshi, "Multiscale pca with application to multivariate statistical process monitoring," *AIChE Journal*, vol. 44, no. 7, pp. 1596–1610, 1998.

[70] P. Nomikos and J. F. MacGregor, "Multivariate spc charts for monitoring batch processes," *Technometrics*, vol. 37, no. 1, pp. 41–59, 1995.

[71] W. Zucchini and I. L. MacDonald, *Hidden Markov models for time series: an introduction using R*. 2009.

[72] B. N. U and L. RAM, "Attribute control charts for markov dependent production processes," 1990.

[73] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd monitoring using image processing," *Electronics & Communication Engineering Journal*, vol. 7, no. 1, pp. 37–47, 1995.

[74] B. E. Moore, S. Ali, R. Mehran, and M. Shah, "Visual crowd surveillance through a hydrodynamics lens," *Communications of the ACM*, vol. 54, no. 12, pp. 64–73, 2011.

[75] D. Helbing, A. Johansson, and H. Z. Al-Abideen, "Dynamics of crowd disasters: An empirical study," *Physical review E*, vol. 75, no. 4, p. 046109, 2007.

[76] U. Soori and M. R. Arshad, "Underwater crowd flow detection using lagrangian dynamics," in *Proceedings of 2nd International Conference on: Underwater System Technology*, 2008.

[77] B. Krausz and C. Bauckhage, "Loveparade 2010: Automatic video analysis of a crowd disaster," *Computer Vision and Image Understanding*, vol. 116, no. 3, pp. 307–319, 2012.

[78] O. Dictionaries, "Oxford dictionaries." Online; accessed 29-April-2015.

[79] T. List, J. Bins, J. Vazquez, and R. B. Fisher, "Performance evaluating the evaluator," in *Proceedings of 2nd Joint IEEE International Workshop on: Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pp. 129–136, 2005.

[80] H. Jhuang, E. Garrote, X. Yu, V. Khilnani, T. Poggio, A. D. Steele, and T. Serre, "Automated home-cage behavioural phenotyping of mice," *Nature Communications*, vol. 1, p. 68, 2010.

[81] E. A. van Dam and L. P. Noldus, "Rat behavior: human versus automatically generated annotation," in *Proceedings of 21st International Conference on: Pattern Recognition (PR)*, 2012.

[82] W. A. Jensen, L. A. Jones-Farmer, C. W. Champ, W. H. Woodall, *et al.*, "Effects of parameter estimation on control chart properties: a literature review," *Journal of Quality Technology*, vol. 38, no. 4, pp. 349–364, 2006.

[83] Matlab, "Computer vision system toolbox(R2012b)," 2012.